Dairybiota: analysing the microbiota of the dairy chain using next generation sequencing

A thesis presented to the National University of Ireland for the degree of

Doctor of Philosophy

By

Conor Doyle B.Sc

Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland

School of Microbiology, University College Cork, Cork, Ireland

2017

Research supervisors: Dr. Paul Cotter and Professor Paul O'Toole

"*And all this science, I don't understand. It's just my job five days a week***"**

Rocket man

"No number of sightings of white swans can prove the theory that all swans are white. The sighting of just one black one may disprove it."

Karl Popper

Table of Contents

Chapter 2: Literature review

Anaerobic sporeformers and their significance with respect to milk and dairy products

Chapter 3:

High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD, is entirely my own work and has not been submitted for another degree, either at University College Cork or elsewhere.

Signed: _____

Conor Doyle

Date:

Thesis Abstract

Advances in DNA sequencing technologies, accompanied with developments in data analysis and interpretation, have provided novel insights into to the microbial ecology of foods and food production environments. By utilising these advances in technology, it is possible to overcome the biases associated with culture-based analysis. This has been achieved by targeting the metagenomic DNA of these environments using high-throughput sequencing (HTS). In this thesis, HTS was utilised to provide insights into factors which influence the microbial composition of raw milk, and reveal potential environmental sources of bacteria in the dairy chain. Firstly, a literature review explores the recent insights gained from applying HTS to study the microbial ecology of food production chains. Additionally, a second literature review focuses on sulphite reducing Clostridia (SRC), their taxonomy, toxigenicity and the prevalence in which they are detected in dairy products. The first research study focused on applying HTS to characterise the microbiota of blended raw bulk tank milk (BTM) stored at different temperatures at both mid and late-lactation. This highlighted that lactation stage had more of a significant impact on the raw milk microbiota compared to storage temperature. After this, in a second study, on-farm environmental niches were explored as possible reservoirs for bacteria to contaminate raw milk. Raw milk samples were collected from individual cows and from BTM, both when cows were housed indoors and when cows were grazing on pasture. Additionally, faecal and teat swab samples were collected from these cows over both periods, as well as environmental niches from both the indoor and outdoor habitats. Results from this study highlight that herd habitat drives the microbial composition of raw milk. In a subsequent investigation, shotgun metagenomic sequencing was employed to explore the cheese production microbiome for the presence of

bacteria and phage, and this approach also facilitated strain level characterisation of starter bacteria. Production plant surfaces were found to harbour resident lactic acid bacteria and brine was identified as a potential reservoir for lactococcal phage. In a final study, whole genome sequencing and *in-silico* genome characterisation were used to determine the genes responsible for the SRC phenotype in dairy associated SRCs. Genome annotation facilitated the identification of two distinct pathways involved in the reduction of sulphite to sulphide in dairy associated isolates, *asrABC* mediated reduction in SRCs and *cysJI* mediated reduction in other sulphite reducing bacteria (SRBs). Ultimately this thesis will show that HTS can be a valuable tool for characterising the microbiota of food products and food production environments.

Publications

Doyle, C. J., D. Gleeson, K. Jordan, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." International journal of food microbiology 197: 77-87.

Doyle, C. J., D. Gleeson, P. W. O'Toole and P. D. Cotter (2017). "High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions." International Journal of Food Microbiology.

Doyle, C. J., D. Gleeson, P. W. O'Toole and P. D. Cotter (2017). "Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study." Applied and environmental microbiology 83(2): e02694-02616.

Doyle, C. J., P. W. O'Toole and P. D. Cotter (2017). "Metagenome-based surveillance and diagnostic approaches to studying the microbial ecology of food production and processing environments." Environmental Microbiology.

Glossary of Terms

^{o}C: Degree's Ceclius

AFLP: Amplified-fragment length polymorphism

ADIM: Average days in milk

asrA: Anaerobic sulphite reductase sub-unit A

asrB: Anaerobic sulphite reductase sub-unit B

asrC: Anaerobic sulphite reductase sub-unit C

BLAST: Basic Local Alignment Search Tool

BAB: Butyric acid bacteria

BAS: Butyric acid spores

Bp: Base pairs

BTM: Bulk tank milk

CAC: Codex Alimentarius Commission

Cfu: Colony forming units

DGGE: Denature gradient gel electrophoresis

DNA: Deoxyribonucleic acid

Ds: double stranded

EFSA: European Food Safety Authority

EMA: Ethidium monoazide

EU: European Union

FAO: Food Agriculture Authroity

FASTA:

FASTQ:

g: G-force

g: Weight in grams

Gb: Gigabase

GFMP: Good fame management practices

HTS: High-throughput sequencing

IB: Infant botulism

ICMSF: International Commission on Microbiological Specification of Foods

INP: Indoor no prep

IP: Indoor prep

ITS: Internal transcribed spacer

L: Litre

Mb: Megabases

MDS: Multidimensional scaling

MDA: Multiple displacement amplification

Mg: Milligram

mL: Millilitre

MPN: Most-probable number

NCBI: National Centre

ng: Nano gram

nM: Nano Molar

ONP: Outdoor no prep

OP: Outdoor prep

OTU: Operational taxonomic unit

P: P value

PCoA: Principal Coordinate

PCR: Polymerase chain reaction

PDO: Product of designated origin

PFGE: Pulse field gel electrophoresis

pH: Power of Hydrogen

PIF: Powdered infant formula

PMA: Propidium monoazide

QIIME: Quantitative insights into microbial ecology

qPCR: Quantitative PCR

RNA: Ribonucleic acid

Rpm: Revolutions per minute

rRNA: Ribosomal RNA

SCC: Somatic cell count

SRA: Short read archive

SRB: Sulphite reducing bacteria

SRC: Sulphite reducing bacteria

TBC: Total bacterial count

TGGE: Temporal gradient gel electrophoresis

VNBC: Viable but not cultivable

List of Figures

List of Tables

Chapter1: Literature Review1

Metagenome-based surveillance and diagnostic approaches to studying the microbial ecology of food production and processing environments

Included as published in Environmental Microbiology (doi:10.1111/1462-2920.13859)

1.0 Abstract

Metagenomic-based analyses have the potential to revolutionise our understanding of the microbiology of food production and processing environments. By adopting such approaches it will be possible to more accurately determine sources of microbial contamination, identify critical control points for such contaminants, and select practices that optimise quality and safety. This mini-review will discuss the merits of adopting metagenostic-based approaches, highlight novel insights that they have provided to date and consider how they could be further implemented.

1.1 Introduction

It has long been recognised that bacteria from food production and processing environments can have positive or negative influences on the end products. Despite the fact that modern food processing facilities are designed to reduce the risk and likelihood of producing spoiled or unsafe produce, they are not abiotic. These facilities are vulnerable to colonisation by microbes from various sources (including raw materials, air, humans and a variety of other sources).

These environments are routinely tested for the presence of pathogenic and spoilage type bacteria, with specific focus on particular species or phenotypes (Doyle, Gleeson et al. 2015). For these assays it is necessary to know in advance what microbe or trait is being assayed. However, applying such targeted approaches means that other microorganisms can escape detection. Indeed, 38.4 million cases of foodborne illness are caused by unidentified microbes in the United States per annum (Scallan, Hoekstra et al. 2011). At present, although there is no data on the volume of food product loss caused by unknown microbes, or not readily culturable microbial agents, there is evidence to suggest that it may be significant (Quigley, O'Sullivan et al. 2016). High throughput DNA sequencing (HTS)-based analysis of metagenomic DNA (DNA from all organisms in an environment) provides a potential means by which the microbiome of sampled environments can be tested to identify unknown, or overlooked, etiological and spoilage agents (Huang, Luo et al. 2016). Here, we outline the benefits of using a microbial ecology-based approach to study the food production and processing facility microbiome and, more specifically, we highlight the advantages of utilising metagenomic-based analyses to further understand these environments. Examples of how these approaches have improved, or potentially will

improve our understanding of microbial influences on some representative food processing/production environments is presented in Figure 1. Ultimately, by adopting these approaches it can be possible to assess how factors such as production practices, building design, seasonality, and operating procedures may be adapted to safeguard the microbial integrity of the food supply chain.

1.2 The microbial ecology of crop production and livestock management

It is important to consider that these approaches are not limited to studying microbial biogeography of the food production facility environment. In order to effectively monitor and control the microbial ecology of food production chains from “farm to fork”, it is necessary to first examine factors which influence the microbiome of crops and animals. In crop production, the rhizosphere (area of microbial rich soil, in immediate contact with plant roots) is an important consideration. In crop production, the host (crop) microbe interplay in the rhizosphere is crucial for plant nutrient acquisition and maintaining crop health (Mendes, Kruijt et al. 2011). Above ground, the phylosphere (bacteria on plant surfaces above ground) may be colonised by potential plant and human pathogens (Rastogi, Coaker et al. 2013). Recent surveillance of the surface microbiota of fresh fruits and vegetables found that farming practice (conventional versus organic) significantly influenced the microbial composition of the food product. These researchers found that *Enterobacteriaceae* were in significantly lower relative abundances in organically farmed

produce (Leff and Fierer 2013). This is notable, as this family contains the genera *Escherichia*, *Shigella* and *Salmonella* which are commonly associated with foodborne illness (Scallan, Hoekstra et al. 2011). However, the amplicon-based approach used for this study was unable to achieve genus, species or strain level classification. This issue regarding discriminatory power will be discussed further below. Nonetheless the study highlights how high-throughput sequencing (HTS) can be utilised to examine the influence of production practices on the fruit and vegetable microbiota. More promisingly, shotgun metagenomic sequencing has recently been used to detect Shiga toxin-producing *Escherichia coli* (STEC) in spiked spinach samples (Leonard, Mammel et al. 2015). Strain level classification was achieved, even at low cell numbers (10 CFU in 100g of spinach), outlining how this technology could be used to conduct culture-independent surveillance of fresh produce for pathogenic microorganism and viruses.

In livestock management, the animal's microbiota is important from the perspectives of diet, nutrient efficiency and animal health (Kim and Isaacson 2015, Schokker, Veninga et al. 2015, Weimer 2015). Compositional metagenomic analysis has been used to analyse the microbiota of bovine teats to identify microbial markers for teat health, which in the future could potentially be used to diagnose mastitis (Falentin, Rault et al. 2016). This type of diagnostic approach has also been used to identity microbial biomarkers associated with Johne's disease, which causes substantial financial losses to farmers whose herds that are affected with this disease (Derakhshani, De Buck et al. 2016). The application of metagenomic analysis has the potential to screen samples of animal origin (herds) for multiple potentially pathogenic microbes in parallel. The most important outcome from applying such an approach would be to maintain animal health and to prevent the transmission of zoonotic diseases to consumers.

In addition to identifying possible biomarkers of disease in livestock, functional metagenomic sequencing has also been utilised to identify antibiotic resistant (AR) genes in bovine manure. This is of importance as manure is frequently applied as a crop fertilizer (Wichmann, Udikovic-Kolic et al. 2014). The use of manure with AR genes could possibly lead to the accumulation of AR genes in the farm environment and the subsequent transmission of microbes with these genes to other environments, animals or to consumers via contaminated agricultural produce.

1.3 Moving inside, biogeography of the food production and processing environment

In recent years, advances in HTS have revolutionised the relatively new field of buildings ecology. This field involves analysis of the biogeography of all fomites (abiotic surfaces that are capable of harbouring microbes) within a given environment and how humans interact with these fomites and the ecosystem as a whole. This is conducted by analysing the influence that, for example, human traffic and extrinsic factors (air flow, temperature and humidity (Kembel, Jones et al. 2012)) have on the distribution of microbes in the environment in its entirety, and more specifically, on niches within these environments. Initial studies in this area focused on hospital and campus buildings (Kembel, Jones et al. 2012, Kembel, Meadow et al. 2014).

Although originally formulated to describe the distribution of microbes throughout the natural environment, the concept on microbial ubiquity propounded by Martinus Beijerinck

that “Everything is everywhere, but the environment selects”, can equally be applied to help understand the distribution of microbes in food processing facilities. For instance, recently it has been shown that niche specific ecosystems have developed within cheese processing facilities. This community development is driven by community-wide adaptation in response to substrates or conditions within each niche (Bokulich and Mills 2013). In this well-established artisanal facility, it was proposed that adapted communities could be considered quasi-domesticated, as they have been selected primarily based on positive attributes that contribute to specific organoleptic characteristics of the cheese (Bokulich and Mills 2013, Bokulich, Ohta et al. 2013). In some incidences, such as *kimoto* rice wine fermentations, the fermentation process is dependent on inoculation with members of the facility microbiome in order to complete the fermentation process (Bokulich, Ohta et al. 2014). The same is also true for production of Ragasano and Salers cheeses, whereby the wooden vats used to store the milk contribute bacteria to the milk to aid in cheese production (Lortal, Di Blasi et al. 2009, Didienne, Defargues et al. 2012). Indeed, for Ragasano cheese, no starter is added, the bacteria which ferment this product are from the wood vat biofilm or the raw milk itself. These examples highlight the beneficial influence that the resident facility microbiome can have on the food production process.

It is important to note that antagonistic microbial adaptations may also be selected for in food processing facilities (Bokulich, Bergsveinson et al. 2015). Indeed, it may be argued that modern food processing facilities contribute in their own way to microbial colonisation due to microbial adaptations; for instance, specific biofilm-forming populations can be selected on stainless steel surfaces that undergo specific cleaning regimes and can become a persistent problem in the dairy processing environment (Sharma and Anand 2002, Cherif-Antar, Moussa–Boudjemâa et al. 2016). This phenomenon is not confined to the dairy

processing environment; recently it was found using HTS that the meat processing environment is home to undesirable microbes that may cause spoilage. In this processing environment, spoilage-associated microbes originate on the carcasses entering the butchering facility (De Filippis, La Storia et al. 2013). Once these microbes have entered these facilities *via* this vector, they establish themselves as resident members of the facility microbiome on the meat contact surfaces. These antagonistic microbes are then subsequently inoculated onto the meat as it is processed into different cuts.

It is evident from the studies highlighted above that the food production facility microbiome can exert a positive or negative influence on the food produced within it. Moving forward, it is reasonable to foresee scenarios where large-scale food producers design facilities to select for microbiomes, or indeed inoculate surfaces with microbes that prevent colonisation by pathogenic or spoilage-associated microbes by way of competitive exclusion (CE). This method has been trialled in a poultry processing plant recently, i.e., drains were treated with CE bacteria (*Lactococcus lactis* subsp. *lactis*), which eliminated detectable *Listeria* in five of the six drains tested (Zhao, Podtburg et al. 2013). In the future metagenomic sequencing could be used to identify additional novel food-grade microbes that already exist within these facilities that could be cultivated on surfaces to confer positives attributes on the production plant ecosystem, mirroring the microbiomes that have evolved in artisanal production facilities.

1.4 Recent advances in understanding the compositional metagenomics of the food production and processing environment

While early approaches to using this technology produced novel and interesting results relating to, for example, milk and cheese, these studies were focused primarily on the microbial composition of the food products and were predominantly curiosity driven. More recently these approaches have become more investigative and hypothesis driven. For instance, they have been used to identify the etiological agent of cheese pinking (Quigley, O'Sullivan et al. 2016). This study identified *Thermus thermophilus* as the causative agent for this type of spoilage (Quigley, O'Sullivan et al. 2016). This bacterium was found in water sources within the production environment, identifying it as a possible reservoir for this contaminant. This is noteworthy because this microbe is not cultivable by any assay currently applied in industry, presumably explaining why it was not previously detected.

The meat processing environment is among the most commonly studied environments. Molecular surveillance of meat processing facilities has been carried out in recent studies in both Finland and Central Europe (Hultman, Rahkila et al. 2015) (Pothakos, Stellato et al. 2015). In both instances, the food spoilage-associated genus *Leuconostoc* was found to be common to the processing environment and the foods; it was highly abundant on raw meat but to be less prevalent on the facility surfaces. Of even greater concern was the prevalence of *Yersinia* spp. on facility fomites among the Finnish processors (who worked in the meat industry); this highlighted that a potentially pathogenic genus was able to survive the cleaning treatments implemented in the processing facility (Hultman, Rahkila et al. 2015). The same genus was, however, found in much lower prevalence in the raw meat products suggesting that the incidence of transmission was limited due to Good Manufacturing Practice. It should also be noted that, while the detection of *Yersinia* spp. is concerning, the identity of the species was not determined, thereby highlighting the importance of using HTS approaches that can assign at the species or even strain level, such as shotgun

metagenomics, in favour of those that only assign at the genus level, such as 16S rRNA amplicon sequencing.

More recently, researchers have demonstrated that there is an observed early variation (significantly different beta diversity) in the microbiota of modified atmosphere packed (MAP) meat between different production lots. This variability in microbial composition is less evident at the end of shelf-life, when beef samples are dominated by *Carnobacterium* spp. and *Brochothrix* spp. (Säde, Penttinen et al. 2017). Bacteria belonging to these groups are among the most frequently associated with meat spoilage and cause the development of off-flavours, slime production, gas production and discolouration (Doulgeraki, Ercolini et al. 2012). The dominance of these microbes was consistent across production lots, with the source identified as the initial meat. Researchers have also compared the influence of different production methods, including different approaches to slaughtering, on the microbiota of meat products. In Halal slaughtered meat *Corynebacteriaceae* was detected in higher proportions than in classically slaughtered meat. Bacterial diversity was also higher in the Halal slaughtered meat (Korsak, Taminiau et al. 2017). The impact of processing on refrigerated pork sausages has also been investigated (Benson, David et al. 2014). In this instance the researchers described the dynamic microbiota of pork sausage throughout its storage, with its initial microbiota being first replaced by *Pseudomonas* spp., and then subsequently by the lactic acid bacteria *Lactobacillus graminis* and *Carnobacterium divergens*.

The dairy production chain, and especially cheese production, has also been the subject of investigations. Recently, researchers in Italy have carried out HTS analysis of a cheese production plant producing two types of cheese (Calasso, Ercolini et al. 2016). They found

that *Streptococcus thermophilus* dominated the fomites of this facility, while several other bacteria that are frequently used as starter cultures were widespread throughout this facility but at much lower levels. *Staphylococcus* and *Brochotrix* were found in both the ripening room and rind of Caciotta cheese, while *Chromohalobacter* and *Sphingomonas* were associated with the ripening room and rind of Caciocavallo Pugliese (Calasso, Ercolini et al. 2016). These results show further evidence of microbial niche partitioning in the food production environment. An additional study which focused on cheese production examined the microbiota of continental style cheese produced at different intervals during the same production day (O'Sullivan, Cotter et al. 2015). This study found that the microbial diversity of cheese samples was higher in cheese produced later in the production day. It was concluded that this was due to the accumulation of bacteria on production surfaces (O'Sullivan, Cotter et al. 2015). HTS surveillance has also been conducted on drain water and drain biofilms in a cheese production plant in Austria (Dzieciol, Schornsteiner et al. 2016). Here, researchers found that the microbiota of the drain water differed from that of the drain biofilm, with *Pseudomonas* spp. being found to be more prevalent in biofilm samples than in the drain water, reflecting the ability of these species to readily form biofilms to survive and persist within production environments. Notably, the ubiquitous foodborne pathogen *L. monocytogenes* was found in the drain water and in the drain biofilm, highlighting both as potential reservoirs for this microbe.

One of the issues to date has been that the vast majority of studies on food processing facilities have been focused on amplicon-based determination of the composition of the microbiota types present. This kind of analysis provides valuable information on the microbial taxa present in any given processing environment and how they may be influenced by different factors, but can only be used to target bacteria or fungi, and offers

limited discriminatory power. This limitation can be overcome by taking alternative approaches; these methods will be discussed in the next section in relation to their application for food safety.

1.5 "Terroir"–ising the microbiome

The idea that microbial communities may imprint distinct organoleptic characteristics upon a food product produced in a facility was touched upon above. The use of HTS has shown that the microbiota of food produce is also influenced by geographical factors including location, culture, and climate. For instance, cheeses produced in geographically distinct regions of the world have distinct microbial communities (Li, Zheng et al. 2017). Bokulich et al. (Bokulich, Amiranashvili et al. 2015) have found that this is also true of another fermented dairy product, matsoni. The results indicated that milk type and production are both drivers of the matsoni microbiota. Indeed, this concept has been examined extensively and first theorised with respect to wine production. The concept of a "Terroir" (the organoleptic signature of a wine, determined by environmental influences) has been explored from a microbial perspective (Bokulich, Thorngate et al. 2014). It was observed that must from different wine producing areas within California have distinct microbiomes. This microbiome is postulated to be shaped by numerous factors including micro-climate, soil type, crop management practices and crop phenotype (Bokulich, Thorngate et al. 2014, Gilbert, van der Lelie et al. 2014).

The microbes associated with production environments shape not only the microbiota of food but also the physical and chemical characteristics of the foods produced within them. Many foods which have Protected Designation of Origin (PDO) have been characterised using metagenomic approaches (De Filippis, La Storia et al. 2014, De Pasquale, Calasso et al.

2014, Dolci, De Filippis et al. 2014, Zinno, Guantario et al. 2017). In the future these methods could be used as a diagnostic tool to prevent producers from selling produce that do not meet the PDO criteria based on microbial composition. Although originally coined to describe a wine organoleptic characteristics based on location and other multifactorial influences, it is now appreciated that the "Terroir" can be expanded to describe other foods (Bokulich, Lewis et al. 2016).

1.6 Food safety

The importance of utilising whole genome sequencing (WGS) of cultured isolates to track specific strains of bacteria involved in outbreaks back to food processing facilities has been reviewed recently (Stasiewicz, den Bakker et al. 2015). WGS is increasingly being successfully applied to trace outbreaks from clinical samples to the source of contamination. The utilisation of WGS in combination with, for example, the Genome Trakr database allows outbreaks to be tracked on global scale (http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm363134.htm).

The successful application of WGS in tracking outbreaks highlights the potential for metagenomic sequencing to bring similar resolution to tracking the movement of microbial communities in the food chain. Indeed, this type of approach has been adopted recently by researchers examining clinical samples from outbreaks of foodborne illness (Huang, Luo et al. 2016). This proof of concept study found *Campylobacter jejuni* in the faeces of a patient suffering from foodborne illness; the pathogen was not present in the patient's faecal

metagenome three months after the incident (when they were healthy). This suggested that the patient was suffering from campylobacteriosis, a fact that is particularly notable given that *Campylobacter* is the most prevalent cause of foodborne illness in the U.S. (Scallan, Hoekstra et al. 2011). It should also be noted that shotgun metagenomics is not limited to just the identification of pathogenic bacterial agents in clinical samples. Indeed, a recent representative study used metagenomic sequencing to identify novel pathogenic agents, such as viruses and parasites, in clinical samples from outbreaks of gastroenteritis of undetermined cause (Moore, Wang et al. 2015). More specifically, the researchers detected viruses, such as rotavirus, adenovirus, sapovirus and parechovirus, as well as the parasite *Dientamoeba fragilis*. However, it is clear that if one is dealing with clinical samples from foodborne outbreaks, then failures have already occurred within the food chain. In order to address this, tracking the movements of microbes in the food chain will be key. This issue is addressed below.

1.7 Microbial Sourcetracking of communities through the food chain

Traditionally, the tracking of microbes through the food chain has been extensively applied to study the movement of pathogenic strains after the occurrence of an outbreak of foodborne illness. While microbial sourcetracking (MST) had originally focused on cultivable microbes such as *Escherichia coli* or *Clostridium perfringens* and took a primarily single-plex approach to tracking these isolates (Scott, Rose et al. 2002), taking culture-independent approaches and targeting nucleic acid make it possible to track contamination in multiplex based assays. This method has been applied in a wide range of studies ranging from the

tracking of species of *Bacteroidales* in water systems (Kapoor, Pitkänen et al. 2015) to the tracking of viruses of human or animal origin in seafood in New Zealand (Wolf, Hewitt et al. 2010).

The development of the SourceTracker algorithm, and its use in combination with high throughput compositional metagenomics has in particular allowed this type of approach to be used to track the movement of microbial communities throughout different environments (Knights, Kuczynski et al. 2011). This Bayesian inference algorithm uses a Markov chain Monte Carlo model to determine potential sources of contamination based on the community composition; "sources" of contamination and "sinks" for contamination. Bokulich and colleagues used the compositional metagenomic data generated from characterising the brewery environment to track the movement of microbes within this environment. They combined this approach with targeted PCR to also track the movement of spoilage-associated genes through this environment. The distribution of bacteria and fungi throughout this environment was also assessed, as were seasonal variations (Bokulich, Bergsveinson et al. 2015). This analysis identified the raw materials used in the brewing process as the main source of bacteria colonising the fomites within this brewery. This approach has also been applied to study microbial movement in the dairy farm environment (Doyle, Gleeson et al. 2016). We found that the raw milk microbiota is influenced by herd habitat and farm management practices (Doyle, Gleeson et al. 2016) i.e. raw milk from herds grazing outdoors had more soil and environmental-type bacteria than raw milk from the same herd when housed indoors during winter, which in turn had higher proportions of gut-type bacteria. This highlighted routes of contamination that need to be managed. (Doyle, Gleeson et al. 2016). A schematic highlighting the different environmental niches that dairy herds are exposed to during different seasons can be seen in Fig.1. Transmission

patterns elucidated from such studies could be used to introduce control measures to reduce or eliminate the risk of transmission in the future.

While SourceTracker provides insightful data which may help elucidate transmission patterns for tracking the movement of bacterial communities from the environment into the food chain, it does not produce information on strain level transmission of microbial movement. The recent development of MetaMLST and Strainphlan has made it possible to track microbial transmission of sequence types and strains of microorganism through environments using metagenomic data sets (Zolfo, Tett et al. 2016). Indeed Strainphlan has been used to track bacterial transmission from mother to infant (Asnicar, Manara et al. 2016). Moving forward, a similar approach to this could be adapted to track etiological agents associated with foodborne illness through the food chain, or to identify critical control points for microbial contaminants within food production environments

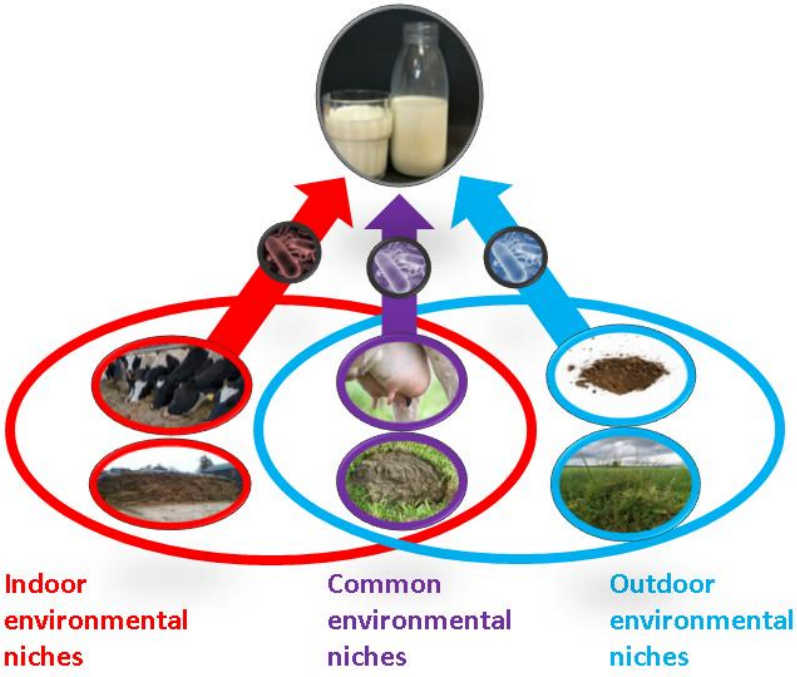


Fig.1: Scematic depicting how environmental niches differ in seasonal milk production. The cow is exposed to different environmental niches when on pasture compared to when grazing. This change in environmental exposure influences the raw milk microbiota (Doyle et al., 2016).

1.8 Application of shotgun metagenomics and metatranscriptomics

Limitations associated with compositional metagenomic surveillance of food production environments, relating to an inability to assign taxonomy at the species level or to provide insights relating to functional potential, can be overcome by utilising shotgun metagenomics. For example, by analysing these shotgun metagenomic datasets it is possible to track movement of bacterial strains or sequence types using the aforementioned MetaMLST or Strainphlan. The advantages of this approach over compositional metagenomics has been outlined in even greater depth recently (Bokulich, Lewis et al. 2016). To date, this type of approach has only been applied to the beef production chain (Yang, Noyes et al. 2016), and has yet to be applied to the food processing environment but has been used to characterise the microbial communities present in fermented foods (Wolfe, Button et al. 2014, Walsh, Crispie et al. 2016) and in cleanrooms (Bashir, Ahmed et al. 2016). In the analysis of the beef production chain (Yang, Noyes et al. 2016), shotgun metagenomics facilitated the detection of pathogens at the species level throughout the production chain. It also allowed for the detection of virulence factors associated with these microbes. Similar methods could be utilised to survey the distribution of microbes throughout the other food processing and production environments.

Furthermore, it is expected that metatranscriptomics will be applied in the future to help characterise microbes present in these environments. Metatranscriptomics is the study of all of the RNA transcribed by a microbial community. In this respect it is similar to metagenomics in that it targets nucleic acid. Unlike metagenomics, which can be used to predict function, metatransciptomics looks at the genes actually expressed by the

community. In a food context, such analyses were recently used to examine the microbial succession in Kimchi fermentations (Jung, Lee et al. 2013) and microbial activity in French cheese (Monnet, Dugat-Bony et al. 2016). In Kimchi fermentations, metatranscriptomic analysis showed that *Leuconostoc mesenteroides* expressed genes involved in the development of flavour at the beginning of the fermentation process. In cheese, Monnet described yeast succession on the surface of Reblachon cheese during cheese ripening phases. The study indicated that these yeast are involved in the production of different flavour compound in the cheese. Two other studies which have utilised a metatranscriptomic approach to food also focus on cheese (Dugat-Bony, Straub et al. 2015, De Filippis, Genovese et al. 2016). Dugat-Bony et al. (Dugat-Bony, Straub et al. 2015) have shown that genes involved in amino acid catabolism are expressed at higher levels in the early phase of cheese ripening, suggesting that this is the most important phase for flavour development. De Filippis and colleagues have showed that ripening temperature could be used to influence the growth of non-starter lactic acid bacteria and that, at the gene level, an increase in temperature saw an up-regulation of genes involved in proteolysis and lipolysis, thereby altering the flavour of the cheese (De Filippis, Genovese et al. 2016). These results show the novel insights that metatranscriptomic analysis can provide with respect to food production in general and, more specifically with respect to these results, it's potential to revolutionise cheese-making through the creation of cheeses with novel flavours and reduced ripening times.

An important consideration when performing metatranscriptomics is the need to deplete ribosomal RNA in order to specifically target the messenger RNA which contains the relevant information highlighting changes in gene expression. Metatranscriptomic analysis, while providing more valuable information than metagenomic analysis, is more expensive

and RNA can be easily degraded making it more difficult to work with. Due to these issues, metatranscriptomics has yet to be extensively applied to studying the food or the food production and processing environments. Its increased application in the future will, however, undoubtedly highlight the components of microbial populations that are more active in individual niches and could, for example, be used to identify genes that confer resilience to microbes that persist within these distinct environments.

1.9 Conclusion

These metagenome-based diagnostic (metagenostic) approaches have already improved our understanding of microbial influences on a few select food production chains (Fig.2). These approaches can be adapted in a similar manner to investigate other food chains. Depending on the diagnostic question being asked, the use of HTS can be tailored to answer any of these questions. Moving forward, it will be necessary to take a more collaborative approach to analyses within these dynamic environments. This will require combining knowledge from both food producers and researchers (from multiple disciplines) for the collective good. By pooling these resources, it will be possible to explore the dynamics of the microbial aspects of our food production in greater detail. Similar collaborative efforts have been established to study the microbiome of other environments, including the human, earth, ocean and hospital microbiome projects to name but a few (Consortium 2012, Smith, Alverdy et al. 2013, Gilbert, Jansson et al. 2014, Sunagawa, Coelho et al. 2015), and have provided intriguing insights. There have been initial efforts to form such consortia, such as the Sequencing of the Food Supply Chain Consortium (SFSCC) (Weimer, Storey et al. 2016).

Ultimately, it is clear that we are only now gaining true insight into the complexity of the food production and processing environments; it is of paramount importance that we use new metagenostic approaches to better design processing facilities and implement control strategies for reducing the ingress of harmful microbes in food production and processing facilities.

Fig.2: Graphic showing how metagen-ostic approaches have and can revolutionise our understanding of the microbial influence on the dairy and wine production scale. Cyan text refers to compositional metagenomics, blue indicates shotgun metagenomics and purple highlights metatranscriptomics.

1.10 References

Asnicar, F., et al. (2016). "Studying vertical microbiome transmission from mothers to infants by strain-level metagenomic profiling." bioRxiv: 081828.

Bashir, M., et al. (2016). "Functional Metagenomics of Spacecraft Assembly Cleanrooms: Presence of Virulence Factors Associated with Human Pathogens." Frontiers in Microbiology **7**.

Benson, A. K., et al. (2014). "Microbial successions are associated with changes in chemical profiles of a model refrigerated fresh pork sausage during an 80-day shelf life study." Applied and environmental microbiology **80**(17): 5178-5194.

Bokulich, N. A., et al. (2015). "Microbial biogeography of the transnational fermented milk matsoni." Food microbiology **50**: 12-19.

Bokulich, N. A., et al. (2015). "Mapping microbial ecosystems and spoilage-gene flow in breweries highlights patterns of contamination and resistance." eLife **4**: e04634.

Bokulich, N. A., et al. (2016). "A new perspective on microbial landscapes within food production." Current opinion in biotechnology **37**: 182-189.

Bokulich, N. A. and D. A. Mills (2013). "Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants." Applied and environmental microbiology **79**(17): 5214-5223.

Bokulich, N. A., et al. (2014). "Indigenous bacteria and fungi drive traditional kimoto sake fermentations." Applied and environmental microbiology **80**(17): 5522-5529.

Bokulich, N. A., et al. (2013). "Monitoring seasonal changes in winery-resident microbiota." PloS one **8**(6): e66437.

Bokulich, N. A., et al. (2014). "Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate." Proceedings of the National Academy of Sciences **111**(1): E139-E148.

Calasso, M., et al. (2016). "Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant." Food microbiology **54**: 115-126.

Cherif-Antar, A., et al. (2016). "Diversity and biofilm-forming capability of bacteria recovered from stainless steel pipes of a milk-processing dairy plant." Dairy Science & Technology **96**(1): 27-38.

Consortium, H. M. P. (2012). "Structure, function and diversity of the healthy human microbiome." Nature **486**(7402): 207-214.

De Filippis, F., et al. (2016). "Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate." Scientific reports **6**.

De Filippis, F., et al. (2014). "A selected core microbiome drives the early stages of three popular Italian cheese manufactures." PloS one **9**(2): e89680.

De Filippis, F., et al. (2013). "Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing." PloS one **8**(7): e70222.

De Pasquale, I., et al. (2014). "Causal relationship between microbial ecology dynamics and proteolysis during manufacture and ripening of protected designation of origin (PDO) cheese Canestrato Pugliese." Applied and environmental microbiology **80**(14): 4085-4094.

Derakhshani, H., et al. (2016). "The features of fecal and ileal mucosa-associated microbiota in dairy calves during early infection with Mycobacterium avium subspecies paratuberculosis." Frontiers in Microbiology **7**: 426.

Didienne, R., et al. (2012). "Characteristics of microbial biofilm on wooden vats (‘gerles’) in PDO Salers cheese." Int J Food Microbiol **156**(2): 91-101.

Dolci, P., et al. (2014). "rRNA-based monitoring of the microbiota involved in Fontina PDO cheese production in relation to different stages of cow lactation." Int J Food Microbiol **185**: 127-135.

Doulgeraki, A. I., et al. (2012). "Spoilage microbiota associated to the storage of raw meat in different conditions." Int J Food Microbiol **157**(2): 130-141.

Doyle, C. J., et al. (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." Int J Food Microbiol **197**: 77-87.

Doyle, C. J., et al. (2016). "High-throughput sequencing highlights the significant influence of seasonal housing and teat preparation on the raw milk microbiota." Applied and environmental microbiology: AEM. 02694-02616.

Dugat-Bony, E., et al. (2015). "Overview of a surface-ripened cheese community functioning by meta-omics analyses." PloS one **10**(4): e0124360.

Dzieciol, M., et al. (2016). "Bacterial diversity of floor drain biofilms and drain waters in a Listeria monocytogenes contaminated food processing environment." Int J Food Microbiol **223**: 33-40.

Falentin, H., et al. (2016). "Bovine Teat Microbiome Analysis Revealed Reduced Alpha Diversity and Significant Changes in Taxonomic Profiles in Quarters with a History of Mastitis." Frontiers in Microbiology **7**.

Gilbert, J. A., et al. (2014). "The Earth Microbiome project: successes and aspirations." BMC biology **12**(1): 69.

Gilbert, J. A., et al. (2014). "Microbial terroir for wine grapes." Proceedings of the National Academy of Sciences **111**(1): 5-6.

Huang, A. D., et al. (2016). "Metagenomics of two severe foodborne outbreaks provides diagnostic signatures and signs of co-infection not attainable by traditional methods." Applied and environmental microbiology: AEM. 02577-02516.

Hultman, J., et al. (2015). "Meat processing plant microbiome and contamination patterns of cold-tolerant bacteria causing food safety and spoilage risks in the manufacture of vacuum-packaged cooked sausages." Applied and environmental microbiology **81**(20): 7088-7097.

Jung, J. Y., et al. (2013). "Metatranscriptomic analysis of lactic acid bacterial gene expression during kimchi fermentation." Int J Food Microbiol **163**(2): 171-179.

Kapoor, V., et al. (2015). "Distribution of human-specific bacteroidales and fecal indicator bacteria in an urban watershed impacted by sewage pollution, determined using RNA-and DNA-based quantitative PCR assays." Applied and environmental microbiology **81**(1): 91-99.

Kembel, S. W., et al. (2012). "Architectural design influences the diversity and structure of the built environment microbiome." The ISME journal **6**(8): 1469-1479.

Kembel, S. W., et al. (2014). "Architectural design drives the biogeography of indoor bacterial communities." PloS one **9**(1): e87093.

Kim, H. B. and R. E. Isaacson (2015). "The pig gut microbial diversity: understanding the pig gut microbial ecology through the next generation high throughput sequencing." Veterinary Microbiology **177**(3): 242-251.

Knights, D., et al. (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature methods **8**(9): 761-763.

Korsak, N., et al. (2017). "Assessment of bacterial superficial contamination in classical or ritually slaughtered cattle using metagenetics and microbiological analysis." Int J Food Microbiol **247**: 79-86.

Leff, J. W. and N. Fierer (2013). "Bacterial communities associated with the surfaces of fresh fruits and vegetables." PloS one **8**(3): e59310.

Leonard, S. R., et al. (2015). "Application of metagenomic sequencing to food safety: detection of Shiga toxin-producing Escherichia coli on fresh bagged spinach." Applied and environmental microbiology **81**(23): 8183-8191.

Li, J., et al. (2017). "Bacterial microbiota of Kazakhstan cheese revealed by single molecule real time (SMRT) sequencing and its comparison with Belgian, Kalmykian and Italian artisanal cheeses." BMC microbiology **17**(1): 13.

Lortal, S., et al. (2009). "Tina wooden vat biofilm: a safe and highly efficient lactic acid bacteria delivering system in PDO Ragusano cheese making." Int J Food Microbiol **132**(1): 1-8.

Mendes, R., et al. (2011). "Deciphering the rhizosphere microbiome for disease-suppressive bacteria." Science **332**(6033): 1097-1100.

Monnet, C., et al. (2016). "Investigation of the activity of the microorganisms in a Reblochon-style cheese by metatranscriptomic analysis." Frontiers in Microbiology **7**.

Moore, N. E., et al. (2015). "Metagenomic analysis of viruses in feces from unsolved outbreaks of gastroenteritis in humans." Journal of clinical microbiology **53**(1): 15-21.

O'Sullivan, D. J., et al. (2015). "Temporal and spatial differences in microbial composition during the manufacture of a Continental-type cheese." Applied and environmental microbiology **81**(7): 2525-2533.

Pothakos, V., et al. (2015). "Processing environment and ingredients are both sources of Leuconostoc gelidum, which emerges as a major spoiler in ready-to-eat meals." Applied and environmental microbiology **81**(10): 3529-3541.

Quigley, L., et al. (2016). "Thermus and the Pink Discoloration Defect in Cheese." mSystems **1**(3): e00023-00016.

Rastogi, G., et al. (2013). "New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches." FEMS microbiology letters **348**(1): 1-10.

Säde, E., et al. (2017). "Exploring lot-to-lot variation in spoilage bacterial communities on commercial modified atmosphere packaged beef." Food microbiology **62**: 147-152.

Scallan, E., et al. (2011). "Foodborne illness acquired in the United States—major pathogens." Emerg Infect Dis **17**(1).

Schokker, D., et al. (2015). "Early life microbial colonization of the gut and intestinal development differ between genetically divergent broiler lines." BMC genomics **16**(1): 418.

Scott, T. M., et al. (2002). "Microbial source tracking: current methodology and future directions." Applied and environmental microbiology **68**(12): 5796-5803.

Sharma, M. and S. Anand (2002). "Characterization of constitutive microflora of biofilms in dairy processing lines." Food microbiology **19**(6): 627-636.

Smith, D., et al. (2013). "The Hospital Microbiome Project: Meeting Report for the 1st Hospital Microbiome Project Workshop on sampling design and building science measurements, Chicago, USA, June 7th–8th 2012." Standards in genomic sciences **8**(1): 112.

Stasiewicz, M. J., et al. (2015). "Genomics tools in microbial food safety." Current Opinion in Food Science **4**: 105-110.

Sunagawa, S., et al. (2015). "Structure and function of the global ocean microbiome." Science **348**(6237): 1261359.

Walsh, A. M., et al. (2016). "Microbial Succession and Flavor Production in the Fermented Dairy Beverage Kefir." mSystems **1**(5): e00052-00016.

Weimer, B., et al. (2016). "Defining the food microbiome for authentication, safety, and process management." IBM Journal of Research and Development **60**(5/6): 1: 1-1: 13.

Weimer, P. J. (2015). "Redundancy, resilience, and host specificity of the ruminal microbiota: implications for engineering improved ruminal fermentations." Frontiers in Microbiology **6**: 296.

Wichmann, F., et al. (2014). "Diverse antibiotic resistance genes in dairy cow manure." MBio **5**(2): e01017-01013.

Wolf, S., et al. (2010). "Viral multiplex quantitative PCR assays for tracking sources of fecal contamination." Applied and environmental microbiology **76**(5): 1388-1394.

Wolfe, B. E., et al. (2014). "Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity." Cell **158**(2): 422-433.

Yang, X., et al. (2016). "Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within the microbiome of the beef production chain." Applied and environmental microbiology **82**(8): 2433-2443.

Zhao, T., et al. (2013). "Reduction by competitive bacteria of Listeria monocytogenes in biofilms and Listeria bacteria in floor drains in a ready-to-eat poultry processing plant." Journal of food protection **76**(4): 601-607.

Zinno, P., et al. (2017). "Impact of NaCl reduction on lactic acid bacteria during fermentation of Nocellara del Belice table olives." Food microbiology **63**: 239-247.

Zolfo, M., et al. (2016). "MetaMLST: multi-locus strain-level bacterial typing from metagenomic samples." Nucleic Acids Research: gkw837.

Chapter 2: Literature review 2

Anaerobic sporeformers and their significance with respect to milk and dairy products

Included as published in the International Journal of Food Microbiology (doi: 10.1016/j.ijfoodmicro.2014.12.022)

2.0 Abstract

Sporeforming bacteria are a significant concern for the international dairy industry. Spores present in milk survive heat treatments and can persist during downstream processing. If they are present in sufficient numbers in dairy products they can cause spoilage or lead to illness as a result of toxin production. While many reviews have highlighted the threat posed by spores of aerobic bacteria to the dairy industry, few have focused on problems caused by the array of different species of anaerobic sporeformers (*Clostridium* and related genera) that can be found in milk. This is despite the fact that members of these bacteria are found throughout the dairy farm environment, and can be toxigenic, neurotoxigenic or spoilage bacteria. This makes the possible presence of *Clostridium* and related spores in bulk tank milk (BTM) important from both a financial and a public health perspective. In this review dairy associated anaerobic sporefromers are assessed from a number of perspectives. This includes the taxonomy of this group of bacteria, the important subgroup of this genus the “sulphite reducing clostridia” (SRC), how these bacteria are detected in milk products, the epidemiological data regarding pathogenic species and strains within the SRC group as well as the influence of farming practices on the presence of SRC in BTM.

2.1 Introduction

Sporeformers are Gram positive bacteria belonging to the phylum Firmicutes. Members of this group form spores when subjected to environmental stresses such as nutrient limitation, osmotic pressure or extreme temperature deviations. These spores, which facilitate survival, are resistant to chemicals (Russell 1990), pH changes (Blocher and Busta 1983), heat, osmotic shock and ultraviolet light penetration (Roberts and Hitchins 1969). When conditions again become suitable for growth, spores can germinate to vegetative cells (Russell 1990). Spores can survive for extended time periods, for example, recoverable spores have been found in dried milk powder from Ernest Shackelton's Cape Royds Hut in Antarctica (Ronimus, Rueckert et al. 2006), and from materials dated to between 25 and 40 million years ago (Cano and Borucki 1995). This robust survival strategy, coupled with the toxigenic potential of some sporeformers, makes sporeforming bacteria a major concern for the food industry(Andersson, Rönner et al. 1995). Spores are frequently associated with silage (Vissers, Te Giffel et al. 2007), soil (Barash, Hsia et al. 2010), forage, animal faeces (Princewell and Agba 1982) and inadequate udder hygiene (Christiansson, Bertilsson et al. 1999), which can in turn lead to their presence in bulk tank milk (BTM). As well as being a concern with respect to raw milk products such as raw milk and artisanal cheeses, the heat stability of spores means that they can also be an issue in commercial dairy products, even when the associated milk has been subjected to heat treatments such as thermisation and pasteurisation (Sugiyama 1951). Indeed, mild heat treatments, such as thermisation, may exacerbate problems by activating spore germination (Griffiths, Phillips et al. 1988, Hanson, Wendorff et al. 2005). Furthermore,

while severe heat treatments such as ultra-high temperature (UHT) and commercial sterilisation are effective at eliminating up to 99.99% of spores (Cox 1975), these heat treatments significantly alter the flavour of liquid milk (Cogan 1977). Their ability to survive exposure to sever heat treatments has led to sporeforming bacteria being referred to as "Thermoduric" bacteria (Gleeson, O'Connell et al. 2013). Processes such as bactofugation can be used to reduce the number of spores and total bacteria in milk. Indeed, this processing step has been demonstrated to achieve a > 95% reduction in total bacterial load (Kosikowski and Fox 1968) and to bring about 60% reduction in spore numbers (Su and Ingham 2000). However, this process is expensive, time consuming and labour intensive (Walstra, Walstra et al. 2010). Microfiltration is another processing step which can be carried out. This process is restricted to skim milk, as spores are roughly the same size as fat globules in whole milk (Rysstad and Kolstad 2006). The requirement for milk fat separation to facilitate microfiltration makes this process labour intensive and expensive to carry out (Skanderby, Westergaard et al. 2009). Ultimately, due to their ubiquitous presence in nature and the frequently high levels at which they are found in particular environmental niches on the dairy farm, it is impossible to eliminate the risk of spore contamination of milk. It is, however, possible to reduce this risk through the implementation of good farm management practices (GFMP) and specific processing steps.

This review will provide an initial overview of the spores of particular importance to the dairy industry before specifically focussing on the importance of anaerobic sporeformers, belonging to the genus *Clostridium*, and, even more specifically, spoilage and pathogenic representatives of this group.

2.2 Different groups of sporeforming bacteria

Sporeformers can be subdivided into different groups based on a number of criteria. These criteria include taxonomy, the specific metabolic capabilities which they possess, their ability to grow at different temperatures or whether or not they can utilise oxygen as a terminal electron acceptor.

2.2.1 Psychrotrophic thermophilic sporeformers

Sporeforming bacteria belonging to the group psychrotrophic thermophilic sporeformers (PTS) are a particular problem to the dairy industry. Members belonging to this group of thermoduric bacteria are able to grow at refrigeration temperature. These PTS typically colonizing raw milk after it is excreted from the mammary gland of lactating cows and multiplying in the bulk tank when the milk is chilled (Murphy, Lynch et al. 1999). Members of the PTS can then survive subsequent heat treatment and processing when in the spore form, and may go on to cause food poisoning, or to limit the shelf life of pasteurised milk and dairy products (Te Giffel, Beumer et al. 1997). Members of the PTS include *Bacillus* species such as *Bacillus cereus*, *Bacillus licheniformis*, *Bacillus subtilis*, *Bacillus amyloliquefaciens* and some members of the *Clostridium* genus (Cousin 1982, Sørhaug and Stepaniak 1997). Two distinct groups of PTS are consistently detected in the dairy production sector i.e. aerobic and anaerobic sporeformers, and an overview of both aerobic and anaerobic PTS is provided below.

2.2.1.1 Aerobic psychrotrophic thermophilic spore formers

Dairy associated aerobic sporeformers belong predominantly to the genus *Bacillus,* with *Paenibacillus* and other genera that were previously assigned to *Bacillus* (Xu and

Côté 2003, Fritze 2004) also being of relevance. Of the *Bacillus spp.* implicated in the contamination of dairy produce, *B. cereus* is considered the most important because of the ability of some strains to induce illness. Toxin producing strains of *B. cereus* can cause two types of food poisoning, i.e. emetic and diarrhoeal. The diarrhoeal toxin is produced as a consequence of spore germination and outgrowth in the small intestine, while the emetic toxin is produced by vegetative cells growing in the milk pre-heat-treatment (Kramer and Gilbert 1989). Recently, it was found that the majority of *B. cereus* group isolates originating from rice were either toxigenic or potentially toxigenic, and that some produced both types of toxin (Oh, Ham et al. 2012). From a spoilage perspective, it is notable that many *Bacillus* sp. can produce thermotolerant lipolytic enzymes which can lead to spoilage of milk. These enzymes exhibit optimum activity at temperatures between 60- 75^{o}C (Schmidt-Dannert, Rua et al. 1997, Chen, Daniel et al. 2003), i.e. temperatures similar to those used for pasteurisation and thermisation. *Paenibacillus sp*. are another group of aerobic bacilli associated primarily with the spoilage of milk and milk products (Huck, Sonnen et al. 2008, Ranieri, Huck et al. 2009). This genus is heavily associated with the spoilage of milk stored in excess of 10 days and, has previously been found to comprise over 95% of the bacterial population present in milk after prolonged refrigeration (Ranieri, Huck et al. 2009, Ranieri, Ivy et al. 2012). *Geobacillus stearothermophilus* (formerly *Bacillus stearothermophilus*) is another aerobic sporeforming species of significance for the dairy industry(Burgess, Lindsay et al. 2010). Together with *Bacillus* spp., *G. stearothermophilis* can cause long term persistent contamination of dairy processing facilities, due to their ability to form biofilms on stainless steel surfaces of processing equipment (Flint, Bremer et al. 1997). It should also be noted that some species

belonging to the genus *Bacillus* and other sporeforming genera are facultative anaerobes. This group includes the most commonly isolated thermophilic sporeforming contaminant in the dairy industry *Amoxybacillus flavithermus* (Ronimus, Rueckert et al. 2006, Burgess, Lindsay et al. 2010). Indeed *B. licheniformis* is often the most frequently isolated mesophilic contaminant in raw milk samples (Waes 1976, Phillips and Griffiths 1986, Crielly, Logan et al. 1994). Some strains of this species have been observed to exhibit accelerated growth in skim milk in an anaerobic environment (Ronimus, Parker et al. 2003).

2.2.1.2 Anaerobic psychrotrophic thermophilic sporeformers

There are also very many anaerobic sporeformers that are problematic for the dairy sector. This group almost is exclusively comprised of current or former members of the genus *Clostridium*, which were first detected in milk and dairy products during the early 20th century (Hussong and Hammer 1930). With respect to refrigeration temperatures, it is notable that some C*. perfringens* strains have a decimal reduction value (time taken at a given temperature to achieve a 90% reduction of vegetative cells) as great as 11 days under standard refrigeration conditions (4°C) (Li and McClane 2006). Furthermore, studies of some *C. botulinum* strains have shown growth and toxin production between 6°C and 8°C (Derman, Lindström et al. 2011) and, indeed, from 3°C to 5°C (Eklund, Wieler et al. 1967, Graham, Mason et al. 1997).

Unlike the aforementioned facultative anaerobes, *Clostridium* spp. are almost all obligate anaerobes and planktonic cells do not tolerate oxygen. However, it has been recently demonstrated that *C. perfringens* can tolerate the presence of oxygen when growing in a biofilm (Charlebois, Jacques et al. 2014). Species belonging to this genus

are, like many other sporeformers, are ubiquitous, being present in soil, in association with nitrogen fixing endophytes on gramineous plant tissue (Minamisawa, Nishioka et al. 2004), in the gastro-intestinal tract of mammals (Brynestad and Granum 2002) and in many other environments. This group of sporeformers, because of their widespread distribution coupled with their spoilage and pathogenic potential are the main focus of the remainder of this review. Before we look at specific spoilage and pathogenic anaerobic sporeformers, it is first necessary to review the taxonomy of the genus *Clostridium.*

2.3 Taxonomy of *Clostridium*

Long before 16S rRNA profiles were applied to define microbial phylogenies, the genus *Clostridium* was defined as containing Gram positive, obligate or strictly anaerobic non-sulphate reducing bacteria (Hippe 1992). It is now apparent that the traditionally classified genus *Clostridium* is very heterogeneous and contains over 100 species, leading to it being divided into distinct clusters based on 16S rRNA gene homology by Collins in 1994 (Collins, Lawson et al. 1994).

Further work has led to the description of a "core" group of species which have been described as "true" *Clostridum* spp.. These species belong to Cluster 1 and Cluster 2 of the clusters established by Collins (Collins, Lawson et al. 1994) and are referred to as *Clostridium sensu stricto* (Wiegel, Tanner et al. 2006). Within the *sensu stricto*, *Clostridium butyricum* is considered to be the cornerstone species, as it was the first to be discovered and classified, by Pasteur, originally having been named "*Vibrion butyrique*" (Durre 2001).

The *sensu stricto* includes the *Clostridium* spp. of most relevance to the food industry and, for this reason, are a particular focus of this review. Notably, the *sensu stricto* group contains the neurotoxogenic species *Clostridium botulinum* and *Clostridium tetani* and the prominent foodborne pathogenic species *Clostridium perfringens*, but the nosocomial pathogen *Clostridium difficile* does not fall within the confines of this core cluster (Wiegel, Tanner et al. 2006). Within the *sensu stricto*, *Clostridium sporogenes* and *C. botulinum* subtype A are closely related and are noticeably divergent from other *C. botulinum* subtypes (Collins, Lawson et al. 1994). The *sensu stricto* also contains a subgroup of spoilage bacteria which are grouped on the basis of a common phenotype and are known as the Butyric Acid bacteria (BAB). These will be discussed in detail below. Another subgroup of anaerobic sporeformers of importance to the food industry is the Sulphite Reducing Clostridia (SRC). While these bacteria have not been defined from a taxonomic perspective, it is likely that many SRC are *sensu stricto Clostridium* spp..

Although there continues to be a reliance on the use of traditional phenotypic assays to detect Clostridia of relevance to the food industry, the DNA-based taxonomic classification of respective *Clostridium* species continues to evolve. In the latest update of Bergey's manual over 50 species previously regarded as belonging to the *Clostridium* genus were reclassified as members of other genera based on 16S rRNA gene homology and physico-chemical properties (Ludwig, Schleifer et al. 2009). This taxonomic relocation was not accompanied with name changes. Some examples include the reclassification of *Clostridium celerecrescens* and *C. difficile* as members of the *Lachnospiraceae* and *Peptostreptococcaceae*, respectively (Yutin and Galperin 2013).

2.4 Sulphite reducing clostridia

The majority of *Clostridium* spp. of relevance to the food industry possess the metabolic ability to reduce sulphite to sulphide under anaerobic conditions to generate cell energy (Wilson and Blair 1924, Prevot 1948, Weenk, Van den Brink et al. 1995). Clostridia with this phenotype are identified as SRCs. It should be noted that there is an abundance of sulphite reducing bacteria that do not belong to the class Clostridia. These bacteria fall within the confines of the taxonomically diverse, yet phenotypically similar (at least with respect to agar-based assays designed to detect SRCs), group of microbes known as the sulphite reducing bacteria (SRBs). SRCs are used as an indicator of faecal or soil contamination, reflecting the aforementioned regular isolation of clostridia from the faeces of warm blooded mammals (Aureli and Franciosa 2002) and from soil (Dodds 1993). Associated agar-based assays rely on this phenotype to test various different food matrices for the presence of SRCs (Gibbs and Freame 1965, Weenk, Fitzmaurice et al. 1991, Prevost, Cayol et al. 2013). These assays rely on the fact that the reduction of iron sulphite to ferrous sulphide by SRCs is accompanied by a black colour change confirming the presence of SRCs. It should be noted, however, that inconsistencies can be observed across the different protocols and microbiological media used (Fuchs and Bonde 1957, Mead 1969). The concentration of sulphite in the agar is an important factor as Weenk found that growth is inhibited at a concentrations above 0.1% but that concentrations of 0.025% are too low for some *Clostridium* spp. to produce black colonies (Weenk, Van den Brink et al. 1995), resulting in a recommendation that a minimum concentration of 0.075% sulphite concentration be used. Furthermore, the presence of glucose may induce gas production, which can make results hard to interpret (Weenk, Van den Brink et al. 1995). The exclusion of

sodium acetate is also recommended by Weenk (Weenk, Van den Brink et al. 1995). This organic acid has been shown to reduce the growth of *C. perfringens* and other *Clostridium* spp. (Weenk, Van den Brink et al. 1995, Juneja and Thippareddi 2004). The current method for the enumeration of SRBs (including SRCs) in food is outlined in the International Organisation of Standards (ISO) document 15213:2003 (Standards 2003). This traditional, phenotype-based assay employed by the dairy industry does not discriminate between species. While it is known that spoilage bacteria such as *C. butyricum*, *Clostridium tyrobutyricum*, *C. sporogenes*, *Clostridium beijernikii* and *C. putrifaciens*, as well as pathogenic species such as *C. perfringens* and *C. botulinum* are SRCs, other species that are not of significance from a spoilage or pathogenic perspective will also be enumerated. Indeed, it has recently been emphasised that there is no statistical relationship between SRC counts and the presence of *C. perfringens* and *C. botulinum* in foods (ICMSF 2014). Another issue to be considered when using this culture based phenotypic enumeration technique is that some strains of facultative anaerobes such as *B. licheniformis* are able to reduce sulphite to sulphide under anaerobic conditions (Weenk, Fitzmaurice et al. 1991). There are also other sulphite reducing sporeforming bacteria from the phylum Firmicutes, that are not clostridia, such as *Desulfotomaculum ruminis* (Campbell and Postgate 1965), a species that has previously been found to be an anaerobic contaminant of processed cheese (Savoy, Font et al. 1981) and *Delsulfotomaculum nigrificans* (Donnelly 1980). Finally, some Gram negative bacteria, such as members of the *Enterobacteriaceae*, are also able to produce sulphide in a sulphite-reductase independent manner, which could also lead to false-positive results (Gibbs and Freame 1965). Following on from these points, it should be noted that an enumeration technique is only considered to be

selective for SRCs if it includes a heat treatment step to eliminate gram negative sulphite reducers and antibiotics to selectively inhibit the growth of other sporeformers, such as *Bacillus* spp., which may also reduce sulphite under anaerobic conditions (Fischer, Zhu et al. 2012). Furthermore, it is understood that there is no all-encompassing agar-based assay for the detection of SRC species, therefore the methodology applied must be carefully selected based on source material and target species (Fischer, Zhu et al. 2012). Thus, it is apparent that there is a need to more clearly establish the identity of the colonies that grow when agar-based SRC assays are carried out and, if it is established that there is a significant percentage of false positives, develop new and more accurate alternatives. The polymerase chain reaction (PCR) has been used for some time to detect clostridia and related genera present in faecal samples (Song, Liu et al. 2004) and, more recently, this approach has been taken to detect *Clostridium* spp. in raw milk (Julien, Dion et al. 2008). In addition to describing the important SRC, *C. perfringens*, below, subsequent sections will also address issues associated with *C. botulinum* and the BAB.

2.4.1 *C. perfringens*

C. perfringens is a food poisoning SRC that can be found in raw milk (McAuley, McMillan et al. 2014). This pathogenic microorganism has a pangenome (the full complement of genes in a species, as species may have large variation in gene content between closely related strains (Medini, Donati et al. 2005, Tettelin, Masignani et al. 2005, Li, Adams et al. 2013, Smokvina, Wels et al. 2013, Hassan, Elbourne et al. 2014)) containing genes encoding at least 17 different toxins and, like many other sporeformers, is ubiquitous in nature (Hatheway 1990). Although the gene for α-toxin,

which is the toxin involved in the development of gas gangrene, is located on the core genome, the majority of genes encoding other toxins involved in human illness are found on mobile genetic elements (Uzal, Freedman et al. 2014). The species has been subdivided into 5 subtypes, named A through E. The typing of strains is determined by the repertoire of major toxins which each individual strain produces (Table 1) (Brynestad and Granum 2002). *C. perfringens* Type A food poisoning is the second most common foodborne illness reported in the United States of America and is estimated to cause 1 million cases annually (Scallan, Hoekstra et al. 2011). *C. perfringens* food poisoning is caused by enterotoxin(CPE) which brings about severe abdominal cramps and diarrhoea (SkjelkvÅLe and Uemura 1977). From an epidemiological perspective, it is worth noting that the *cpe* gene can be either chromosomally or plasmid encoded in type A strains. This is important because these strains are considered distinct from one another (Lindström, Heikinheimo et al. 2011). *C. perfringens* isolates with chromosomally encoded *cpe* genes form a homogeneous group, while isolates with plasmid encoded *cpe* genes are heterogeneous and group closely with other cpe-negative *C. perfringens* type strains (Lindström, Heikinheimo et al. 2011). It was previously understood that type A *C. perfringens* food poisoning was caused only by strains with chromosomally encoded *cpe* due to these strains being more resistant to heating, osmotic shock and low temperatures (Sarker, Shivers et al. 2000, Li and McClane 2006, Li and McClane 2006). However, it is now appreciated that strains with plasmid encoded *cpe* can also cause foodborne disease (Lahti, Heikinheimo et al. 2008, Lindström, Heikinheimo et al. 2011). Indeed, strains with plasmid encoded *cpe* have recently been found to exceed the growth potential of the more robust strains with

Table 1: Toxins produced by C. perfringens by strain adapted from Stiles et al., 2013.

Toxin	Subtype				
	A	**B**	**C**	**D**	**E**
α	✓	✓	✓	✓	✓
β		✓	✓		
ε		✓		✓	
ι					✓

chromosomally encoded *cpe* at 12°C (Xiao, Wagendorp et al. 2014). This trait is of significance for raw foods, or any foods which have been contaminated post processing (Xiao, Wagendorp et al. 2014). *C. perfringens* type A strains have also been demonstrated to produce a β2 toxin (Gibert, Jolivet-Renaud et al. 1997). This toxin is present in a low number of type A strains with chromosomally encoded *cpe*, but appears to be prevalent in the majority of strains with plasmid encoded *cpe* (Fisher, Miyamoto et al. 2005). This toxin is postulated to act as an accessory toxin, increasing the ability of CPE to induce illness (Fisher, Miyamoto et al. 2005). The β2 toxin is also produced by type B, C and D strains of *C. perfringens* (Gkiourtzidis, Frey et al. 2001, Jabbari, Tekyei et al. 2012). While Type C *C. perfringens* associated food poisoning is much rarer, it attracts attention as it can be responsible for necrotising enterocoloitis (NEC). This illness is a result of infection with β-toxin producing strains of *C. perfringens*, which usually also produce δ-toxin and θ-toxin, all of which are produced during vegetative cell growth (Brynestad and Granum 2002). However, under normal conditions β-toxin is susceptible to cleavage by trypsin, and so disease can only ensue in situations of trypsin inhibition (Gui, Subramony et al. 2002). Individuals with type I and type II diabetes mellitus are at risk of developing NEC if they consume contaminated foodstuffs because of reduced trypsin activity (Gui, Subramony et al. 2002). The θ–toxin is a cytolytic/haemolytic enzyme, and δ-toxin is also a haemolysin (Cavalcanti, Porto et al. 2004). The largest dairy associated *C. perfringens* outbreak occurred in the United Kingdom in 1981, with 77 school children suffering from *C. perfringens* food poisoning due to the consumption of contaminated milkshakes (Anon 1982). No serological analysis was carried out on the responsible strain, so subtype of *C. perfringens* or toxin produced was never identified. More recently, studies have

detected the presence of *C. perfringens* in infant formula (Barash, Hsia et al. 2010). However no case of *C. perfringens* illness has ever been attributed to the consumption of powdered infant formula.

As well as causing human illness, specific subtypes of *C. perfringens* have also been repeatedly identified as the etiological agents of bovine mastitis (Ribeiro, Lara et al. 2007, Osman, El-Enbaawy et al. 2009) and mastitis in other ruminants (McDonnell and Holmes 1990). Indeed C*. perfringens* had been identified as a causative agent of intra mammary infection as far back as 1977, when it was known as *Clostridium welchii* (Robinson and Manser 1977). Subtype B, D and E strains have all also been associated with the onset of disease in ruminants resulting from toxin production (Stiles, Barth et al. 2013). As they have never been reported to be involved in human illness, these subtypes are not regarded as playing a major role in human illness. An overview of the subtypes involved in both human and animal disease is provided in **(Table 2)**. Specific enumeration techniques do exist for the detection of *C. perfringens* in foodstuffs(Fischer, Zhu et al. 2012). The most commonly applied methods use trypotose-sulphite-cycloserine (TSC)-based agars as referred to in ISO document 7937:2004 (Standards 2004) or, more recently, chromogenic agar-based enumeration techniques have been employed to detect the presence of this pathogen in food (Manafi 2000).

Table 2: Overview of the diseases cause by subtypes of C. perfringens adapted from Stiles et al., 2013.

Subtype	Disease caused
A	Myonecrosis (gas gangrene in humans and animals) NEC in fowls and piglets, food poisoning in humans
B	Hemorrhagic enteritis in calves, foals and sheep, dysentery in lambs
C	NEC in humans, young ruminants and foals, enteroxemia in sheep
D	Enterotoxemia in lambs, goats and cattle
E	Enterotoxemia in calves and lambs

2.5 *C. botulinum*

Due to its neurotoxigenicity, coupled with the frequent isolation of *Clostridium* spp. from on farm environments, *C. botulinum* and the threat posed by this species will constitute an entire section of this review. Although rare, the potential presence of *C. botulinum* in powdered dairy products is a significant concern for the dairy industry and consumers. *C. botulinum* produces a highly toxigenic neurotoxin, often regarded as one of the most potent toxins known to man (Dhaked, Singh et al. 2010), which can induce neuroparalytic disease (Sobel 2005). Strains of *C. botulinum* are classified based on the type of neurotoxin that they produce and, until recently there were seven recognised subtypes (A-G) of this species, which produced seven different variants of the botulinum toxin (Mahant, Clouston et al. 2000, Ting and Freiman 2004). Some strains possess two types of toxin-encoding genes, and are described as being of subtype Ab, Ba, Af, Bf etc., with the uppercase character indicating the more highly expressed toxin gene (Dover, Barash et al. 2014). Indeed, one strain possessing genes encoding three types of toxins has also been described (Gimenez and Ciccarelli 1978, Hatheway and McCroskey 1987, Santos-Buelga, Collins et al. 1998, Barash and Arnon 2004, Dover, Barash et al. 2014). Recently, an eight type of botulinum toxin type H has been associated with a strain isolated from the faeces of an infant suffering from infant botulism (IB) (Dover, Barash et al. 2014). Interestingly, *C. botulinum* is similar to *B. cereus* in its epidemiology, as it can cause two types of foodborne illness, intoxication (ingestion of toxin) and toxicoinfection (ingestion of spores followed by subsequent germination and toxin production *in situ*) (Martin 2003).

The detection of *C. botulinum* in food is difficult to confirm, due to the high degree of genetic homology shared between proteolytic strains of this pathogenic species and the non-toxogenic species *C. sporogenes*. Indeed, this close relationship has previously lead to an incorrectly reported incident of *C. botulinum* contamination in New Zealand in 2013 (Doyle and Glass 2013). It is also worth noting that, according to current classification criteria, some strains of *C. botulinum* would be considered sufficiently distinct to be classified as distinct species (Peck, Plowman et al. 2010). Furthermore, some proteolytic strains of *C. butyricum* and *C. baratii* can also produce botulinum toxin E and F (Hall, McCroskey et al. 1985). Ultimately, while the sequence similarity shared between *C. sporogenes* and *C. botulinum*, coupled with the ability of other *Clostridium* species to produce botulinum toxins make monitoring and confirming cases of foodborne botulism by traditional microbiological techniques difficult, the application of multiplex PCR based assays has the potential to address this issue (Lindström, Keto et al. 2001). With respect to this potential, it should be noted that PCR cannot distinguish between DNA originating from live cells and that from dead cells (Josephson, Gerba et al. 1993) unless some intervention is employed to inactivate the latter (Nocker, Sossa-Fernandez et al. 2007). Extracting DNA from spores is also more challenging than extractions from vegetative cells and, so, unless specific steps are taken to overcome this problem, can lead to an underestimation of *Clostridium* numbers.

2.5.1 Botulism and dairy powders

Human botulism is typically caused by *C. botulinum* strains producing type A, B, D and E toxins (Shapiro, Hatheway et al. 1998), while strains producing type D and C toxins

are most heavily associated with animal botulism (Prévot, Tweepenninckx et al. 2007, Nakamura, Kohda et al. 2010). Type A botulinum toxin is used in the cosmetics industry and marketed under the name “Botox”. Spores of *C. botulinum* can be consumed by the majority of individuals without prompting any illness (Sobel 2005). However, infants between the ages of 2 to 32 weeks and immunocomprimised individuals with an unhealthy intestinal microbiota are at risk of developing disease following from consumption of food contaminated with spores producing type A toxin (Doyle and Glass 2013) .

Infant botulism is a type of toxicoinfection in which spores of *C. botulinum* are ingested by infants. This ingestion coupled with the underdeveloped intestinal microbiota and high pH of the infants gut provides an ideal environment for *Clostridium* spore germination and subsequent toxin production (Arnon 1980). The infectious dose of *C. botulinum* is unknown, but has previously been estimated to be within the range of 10-100 spores (Arnon, Midura et al. 1979). Since the first reported case of IB in 1976 (Pickett, Berg et al. 1976), over 1500 cases have been reported in the United States alone (Brook 2007). The majority of these cases have been attributed to the consumption of contaminated honey or to environmental sources (Brook 2007). In the United States, IB has been particularly prevalent in the state of California (Johnson, Tepp et al. 2005), where the high occurrence can be accounted for by the high birth rate of that state (Arnon 1998). Other countries in which cases have been reported include, Australia (May, Coulthard et al. 2002), Italy (Fenicia, Da Dalt et al. 2002), Denmark (Balslev, Østergaard et al. 1997) and Japan (Kakinuma, Maruyama et al. 1996). The majority of reported cases have been caused by strains producing type A or B botulinum toxin (Brook 2007), with a smaller subset of cases involving type E toxin

producing *C. butyricum* strains (Suen, Hatheway et al. 1988) and type F producing strains of *C. baratii* (Hall, McCroskey et al. 1985, Suen, Hatheway et al. 1988). Although it has long been suggested that a powdered infant formula (PIF) may act as a vehicle for transmission of Clostridia (Brett, McLauchlin et al. 2005, Johnson, Tepp et al. 2005), and indeed *Clostridium* species have been frequently isolated from infant formula and from other dairy powders (Barash, Hsia et al. 2010), no case of IB has ever been definitely attributed to contaminated powders (Johnson, Tepp et al. 2005). This has led WHO/FAO (World Health Organisation/Food and Agriculture Organization) to classify members of this genus as category "C" organisms (causality less plausible or not yet demonstrated to have caused an outbreak) with respect to infant formula microbiological risk assessment (FAO/WHO 2004). Despite this, there has been some concern in that two studies examining samples from the sixth reported case of infant botulism in the United Kingdom in 2001 reported the presence *C. botulinum* spores in infant formula (Brett, McLauchlin et al. 2005, Johnson, Tepp et al. 2005). Both studies involved a molecular analysis, i.e. pulse field gel electrophoresis (PFGE) (Johnson, Tepp et al. 2005) and amplified-fragment length polymorphism (AFLP) (Brett, McLauchlin et al. 2005) respectively, of isolates from samples taken from the infants home, including an open and an unopened can of PIF and of the infant's faeces. The AFLP analysis of the isolates from powder samples taken from the open container revealed four AFLP patterns, of which two patterns were identical to two patterns generated from AFLP analysis of isolates from clinical samples (Brett, McLauchlin et al. 2005). Similarly, the PFGE results found that 1 PFGE profile of an isolate from the opened can of PIF matched a PFGE profile of an isolate from the clinical sample (Johnson, Tepp et al.

2005). Thus it was concluded that the open can of PIF was contaminated from an unknown environmental source and was responsible for the illness.

In light of the reported case of IB in the United Kingdom in 2001, and the debate surrounding its presumptive link to the ingestion of PIF, Barash et al (Barash, Hsia et al. 2010) carried out a two year surveillance of PIF in the state of California. In that study, samples of PIF were obtained from the families of infants being treated for IB while others were purchased from retail outlets. Samples were grouped according to producer and tested for the presence of *Clostridium* species. Attempts were also made to purify botulinum toxin and mouse bioassays were carried out in instances where there was some suspicion of the presence of toxin. The mouse bioassay was negative for all isolates tested, showing that the PIF samples tested were negative for neurotoxigenic *Clostridium* spp.. While no botulinum-producing *Clostridium* was detected, many other soil dwelling *Clostridium* species were found in the samples tested, including, pathogenic species such as *C. perfringens*, *Clostridium septicum*, C*lostridium bifermentans* and *Clostridium novyi* as well as *C. sporogenes*. While the pathogen *C. perfringens* has already been referred to above, the presence of *C. septicum* is notable in that the species has previously been identified as causing myonecrosis in infants with congenital neutropenia (Barnes, Gerstle et al. 2004), and *C. bifermentans* has previously been isolated in a mixed infection with *Peptostreptococcus* sp. in a case of paediatric infection (Brook 1995). Also present in some of these samples was the spoilage associated sporeformer *C. tyrobutyricum.*

2.5.2 Botulism and other dairy products

Although rare (i.e. below 1% of the total incidences of reported foodborne botulism outbreaks), outbreaks of botulism have previously resulted from the consumption of contaminated cheese and other dairy produce (Collins-Thompson and Wood 1992). One of these rare cases involved a jarred cheese product which contained 50 Mouse Lethal Doses (MLD) of type B toxin per gram of cheese. Consumption of 70g of contaminated product resulted, 3 days later, in the death of one individual (Meyer and Eddie 1951). While other samples from the same batch were found to be contaminated with *C. botulinum* type B toxin-producing strain, no toxin production was detected, suggesting that the case of botulism resulted from incorrect storage of this product. Two outbreaks of botulism were reported to be the result of consumption of Brie ripened cheese in France and in Switzerland in 1973, arising due to the manner in which the cheese was stored during the ripening process. During this process cheeses were stored on straw contaminated with animal faeces (Sebald, Jouglard et al. 1974, Billon, Guerin et al. 1980). An outbreak of botulism in the United Kingdom in 1989 resulted from the consumption of yoghurt containing type B botulinum toxin was reported (Critchley, Hayes et al. 1989), 27 cases patients were identified, with 12 admitted into intensive care and one patient dying. The source was identified as hazelnut conserve that was used to flavour the yoghurt but which had not undergone sufficient heat sterilisation (Critchley, Hayes et al. 1989). The toxin concentration in the yoghurt was found to be within the range of 14-30 MLD/mL (O'Mahony, Mitchell et al. 1990). In another instance, an outbreak of botulism in southern Italy, in which 8 people became ill after consumption of mascarpone or tiramisù (mascarpone base desert), was attributed to a type A strain of *C. botulinum* (Aureli, Franciosa et al. 1996). As a

result of this outbreak a surveillance of mascarpone cheese in Italy was undertaken, which found that almost one third of samples were positive for the presence of botulinum spores. The majority of isolates were also type A, with the remaining isolates identified as *C. botulinum* type B. It was concluded that the unusually high prevalence of botulinum spores in mascarpone could be attributed to both the high pH of the final product and ineffective processing (pasteurisation and ultrafiltration) and storage practices (Franciosa, Pourshaban et al. 1999). Finally, a French study published in 2004 focused on the presence of *C. botulinum* in raw ingredients used to manufacture processed foods, which are subjected to long term storage under refrigeration temperatures. It reported that over 10% of dehydrated dairy ingredients tested positive for the presence of *C. botulinum.* The analysis was carried out using PCR in combination with enzyme linked immunosorbant assay (ELISA) (Carlin, Broussolle et al. 2004).

2.5.3 Regulations

A recent report published by the International Commission on Microbiological Specification of Foods (ICMSF) on testing dairy powders for the presence of *C. botulinum* spores noted that at present it is not practical to test specifically for *C. botulinum* and botulinum toxin producing strains of *Clostridium* (ICMSF 2014). Accompanying this report was a publication by Codex Alimentarius Commission (CAC 2008), which concluded that *C. botulinum* was not to be considered a significant hazard in infant formula. The ICMSF suggested to instead test for the presence of SRCs. As noted above, the presence of this group of *Clostridium*, are indicative of soil and/or faecal contamination (ICMSF 2014). Regulations relating to the presence of

Clostridium, or more specifically SRCs, vary, for example the Russian Federation has set a regulatory limit of between 25-100 cfu/g of dairy powder (Federation 2008, ICMSF 2014), while in the United States of America an advisory maximum level of SRCs has been set at between 10-25 cfu/g (Russell 1990, Council 2013). Currently there are no specific regulations in the European Union regarding the presence of *Clostridium* spores in milk or milk products. However, many milk processors have introduced milk quality payment schemes whereby payments to farmers are based on how their milk scores on a number of criteria such as somatic cell count (SCC), butyric acid spores (BAS), TBC, SRCs, presence of antibiotic residues and milk cleanliness (Velthuis and van Asseldonk 2010). These regulations and payment schemes incentivise dairy farmers to adhere to GFMP when producing milk. From a consumer's perspective, it is desirable to have a low tolerance for SRCs in food products, extending to a zero tolerance for neurotoxigenic species such as *C. botulinum*, however despite this, strict regulatory guidelines have yet to be applied globally.

2.6 Butyric acid bacteria

The BAB are a group of mainly spoilage bacteria that also belong to the genus *Clostridium,* and have been referred to briefly above. This group of microbes is associated with the spoilage, through gas defects, of continental cheeses as a consequence of the fermentation of lactate to acetate, butyrate and hydrogen gas (Klijn, Nieuwenhof et al. 1995). Interestingly, other C*lostridium* spp. are the causative agents of "blown pack" spoilage of raw refrigerated vacuum packed meats (Moschonas, Bolton et al. 2010, Silva, Paulo et al. 2011). These *Clostridium* spp. are not

BAB, and thus are not of concern to the dairy industry and, therefore, will not be discussed further. *Clostridium* species belonging to this group are also SRCs, and include the species *C. butyricum*, *C. tyrobutyricum,* and *C. beijerinckii.* The ability of some strains of C*. butyricum* to produce botulinum toxin makes the detection of this group of microbes in milk very important.

BAB contamination of milk and subsequently cheese has been attributed to dairy cow consumption of poor quality silage which has undergone aerobic deterioration, leading to insufficient acidification, and in turn, allowing for *Clostridium* spore germination and growth (Pahlow, Muck et al. 2003). High numbers of BAB spores have been isolated from grass and alfalfa silage, with low spore counts being observed in corn silage (Stadhouders and Jørgensen 1990). It has been postulated that this is because grass silage is more likely to be contaminated with faeces than silage fermented from other substrates. Moreover, spore counts from silage originating from manure spread grass have been proven to be greater than those from grass fertilized with chemical fertilizer (Rammer 1996, Te Giffel, Wagendorp et al. 2002). Thus, the control of BAB spore counts in silage is necessary to in turn prevent the surface contamination of teats and limit BAB spore counts in BTM (Te Giffel, Wagendorp et al. 2002, Vissers, Driehuis et al. 2006).

2.6.1 Gas defects in cheeses

As mentioned above, the presence of BAB spores in raw milk can cause the development of gas defects such as “late blowing” in cheeses such as Gouda , Comté, Emmental and Beufort (Le Bourhis, Doré et al. 2007). *C. butyricum*, *C. tyrobutyricum* and *C. beijerinckii* are associated with these defects (Vissers, Driehuis et al. 2006). Gas

defects may also be caused by *C. sporogenes* which, although not considered a BAB, it can produce gas due to proteolysis in the anaerobic cheese environment (Goudkov and Sharpe 1965, Ting and Freiman 2004, Le Bourhis, Doré et al. 2007). Incidences of butyric acid spoilage of cheese lead to considerable loss of product value. BAB, while all members of *Clostridium sensu stricto*, do not all use the same substrates in their catabolism of the cheese. While *C. sporogenes* may ferment lactate *via* the Embden-Meyerhof-Parnas pathway, it has a substrate preference for amino acids which it metabolises by Strickland reactions (Cato, George et al. 1986, Allison and Macfarlane 1990, Le Bourhis, Doré et al. 2007). *C. beijeinickii* and *C. tyrobutyricum* both utilise the Embden-Meyerhof-Parnas fermentation pathway under optimal conditions, but *C. beijeinickii* can adapt to a solventogenic catabolism under certain pH conditions (Yan, Zhu et al. 1988, Rogers and Gottschalk 1993, Le Bourhis, Doré et al. 2007). It was also proposed by Bourhis (Le Bourhis, Doré et al. 2007) that the metabolites produced by *C. beijeinickii* and *C. sporogenes* may stimulate the growth of *C. tyrobutyricum.*

Traditionally BAB spore numbers were estimated by the most-probable number (MPN) assay based on gas production in anaerobically incubated samples in liquid media (Le Bourhis, Doré et al. 2007). However, it is now recognised that this method and other culture dependent methods are too time consuming and labour intensive. Advances in nucleic acid technology have facilitated the sequencing of *Clostridium* species genomes and have enabled the design of primers to detect spoilage *Clostridium* spp. in raw milk. Early PCR based studies identified *C. tyrobutyricum* as the BAB most frequently associated with gas defects in cheese (Klijn, Nieuwenhof et al. 1995). More recently, gradient gel electrophoresis based approaches, such as denature gradient gel electrophoresis (DGGE) and temporal gradient gel electrophoresis (TGGE) have been

used to study these populations of spoilage clostridia. In one such study, DGGE (coupled with traditional plating approaches) was used to study the microbial population in abnormally ripened/spoiled Grana Pando cheese. Based on the DGGE migrations patterns observed, the majority of the defective cheeses contained only one associated spoilage species. While this sole spoilage agent was most frequently *C. tyrobutyricum*, *C. beijerinckii*, *C. sporogenes* and *C. butyricum* were also identified in other instances (Cocolin, Innocente et al. 2004). Subsequently, TGGE was used to differentiate between the spoilage-associated species of *Clostridium*, being found to have a minimum detection limit of 100 CFU/g (Le Bourhis, Doré et al. 2007). It is also noteworthy that a large scale study of the microorganisms associated with gas defects in cheese made from ovine milk led to the isolation of 233 *Clostridium* isolates from 45 defective Manchego. Through PFGE analysis a number of distinct pulsotypes were identified which grouped together according to the factory in which they were produced and their date of manufacture. It was also observed that some pulsotypes were particularly associated with cheese that had undergone severe late blowing defect (Garde, Gaya et al. 2012). It should be noted, however, that the labour intensive nature of these, and other, electrophoresis-based approaches limits their application from an industrial perspective. More notably, highly specific real-time PCR based assays, such as that developed to specifically detect *C. tyrobutyricum* (López-Enríquez, Rodríguez-Lázaro et al. 2007), or multiplex PCR based assays, such as that developed to targeting all known members of the BAB (Cremonesi, Vanoni et al. 2012), may in the future be adapted for use in an industrial setting.

2.7 Importance of Farming Practices

The application of GFMP is critical to achieving low spore contamination of raw milk. While the dairy industry relies on pasteurisation to achieve a reduction in the number of pathogenic and spoilage microorganism, pasteurisation is ineffective against spores (Gleeson, O'Connell et al. 2013).

As specified earlier, silage type and quality have a considerable impact on the presence of thermoduric bacteria in BTM, with high numbers of both aerobic and anaerobic bacteria having previously been associated with poor quality silage (Vissers, Driehuis et al. 2007, Vissers, Driehuis et al. 2007, Julien, Dion et al. 2008, Garde, Arias et al. 2011). Factors which influence silage quality include starting material, fermentation conditions, pH achieved, dry matter content and contamination level (Rammer 1996, Vissers, Driehuis et al. 2007). Herd consumption of poor quality silage, followed by the survival of spores in the gastrointestinal tract and the contaminating of manure, can subsequently result in the adhesion of contaminated faeces to teats and udder surfaces causing contamination of raw milk as a result (Bergère, Gouet et al. 1968). When cows are on pasture teats can also be contaminated, as the soil microecosystem is abundant in sporeforming bacteria, and particularly those belonging to the genus *Clostridium* (Slaghuis, Te Giffel et al. 1997, Christiansson, Bertilsson et al. 1999). Using predictive microbiology and applying a probability model, it has been estimated that when teats are contaminated with soil, one third of BTM will contain over 1,000 spores per litre compared to a probability of only 2% of BTM containing the same concentration of spores if the contamination of teats was feed related (Vissers, Te Giffel et al. 2007). Udders and teats may also be contaminated with undesirable microorganisms from poor quality and contaminated bedding material. This source of

contamination is typically a problem during the winter months when herds are housed indoors (Magnusson, Christiansson et al. 2007). The bedding material which is commonly used in housed cubicles by farmers is sawdust, but it is now recognised as a reservoir for sporeforming bacteria (Magnusson 2007).With respect to sheep, a recently published study examined the influence of farming practices on the presence of lactate fermenting *Clostridium* spp., i.e. BAB, spores in ewes milk and cheese (Arias, Oliete et al. 2013). It was found that the risk of milk contamination with $>10^3$ spores/ L was almost two and a half times greater if herds were fed on-farm prepared total mixed ration when compared to herd that were fed commercial total mixed ration (Arias, Oliete et al. 2013). Similarly, it was also calculated that feeding wet brewers grains (a cheap by-product from the brewing industry) instead of commercial total mixed ration, increased the likelihood that BTM contained $>10^3$ spores/L by almost four times. In-parlour practices such as dipping teats in cleaning agents pre- and post-milking, have been shown to reduce the bacterial load cows teats and subsequently of BTM (McKinnon and Pettipher 1983, Stadhouders and Jørgensen 1990). While the cleaning agents used in these practices do not destroy spores, when coupled with teat drying with individual wipes prior to cluster attachment, they can reduce the incidence of thermoduric contamination (Jayarao, Pillai et al. 2004, Gleeson, O'Connell et al. 2013), and lower the numbers of *Clostridium* spores in the raw product (Stadhouders and Jørgensen 1990). Indeed, using *C. tyrobutyricum* as a model contamination organism, it has been shown experimentally that adherence to good teat cleaning can lead to a substantial decrease in spore contamination (Melin, Wiktorsson et al. 2002). With regards to cleaning and drying of teats, the type of drying material chosen can significantly affect the cleaning efficiency (Magnusson, Christiansson et al. 2006). A

good parlour cleaning routine after milking is also important with respect to removing dirt and faeces remaining after the herds transition through the parlour (Dodds 1993, Te Giffel, Beumer et al. 1995). Just as for teat preparation and cleaning, the selection of cleaning agent and the concentration thereof for cleaning of milking equipment and machinery is also critical to reducing the risk of milk contamination (Murphy and Boor 2000, Reinemann, Wolters et al. 2003). The water temperature used for cleaning is also critical to reducing the probability of high counts of thermoduric bacteria in BTM (Reinemann, Wolters et al. 2003, Elmoslemany, Keefe et al. 2009). With respect to BTM itself, it is crucial to achieve rapid BTM cooling to 4^{o} C or below within a half hour of the conclusion of milking (Gleeson, O'Connell et al. 2013). The combination of an effective parlour cleaning regime coupled with udder cleaning can also prevent the development and spread of mastitis (Pankey 1989, Schreiner and Ruegg 2003). Ultimately, adherence to GFMP is necessary in parlour, on pasture and in-house to reduce the risk of spores contaminating BTM and to maintain herd health.

Finally, the resilience of clostridial spores in the dairy farm environment can perhaps be explained best by the concept of "the clostridial spore contamination cycle" put forward by Pahlow (Figure.1) (Pahlow, Muck et al. 2003). This describes how contamination with spores can originate from the soil environment and from organic fertiliser residues during silage harvesting. This, combined with favourable spore germination conditions during subsequent silage fermentation, can lead to an increase in spore numbers in the silage. Contaminated silage is then consumed by cows and spores survive the alimentary transit and accumulate in excrement. This waste may contaminate cow's teats and cause bulk milk contamination during milking. Likewise, contaminated faeces may also be released back into the soil when organic fertiliser

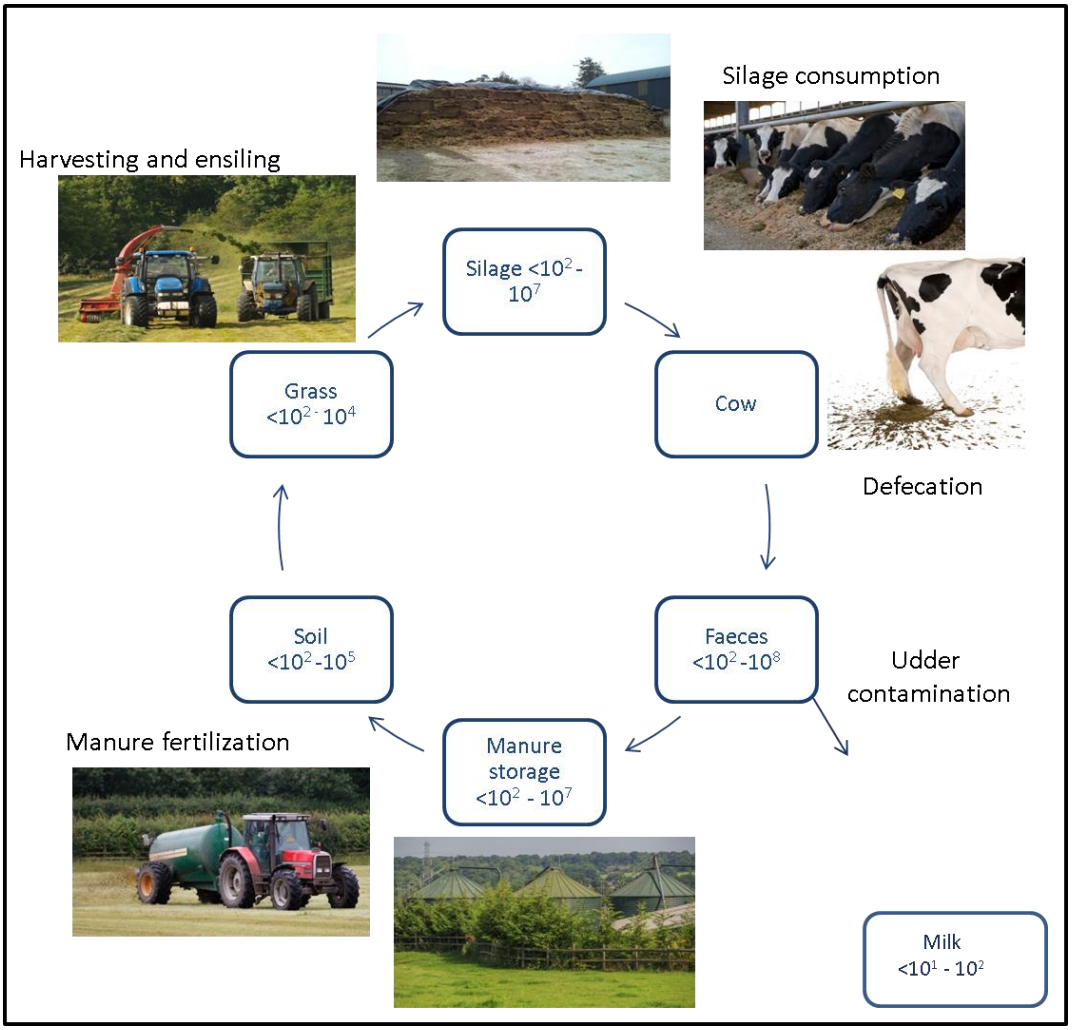


Fig.1: Contamination cycle of Clostridium spores on dairy farms. Adapted from Pahlow et al. (2003).

(manure) is spread on land. Thus, clostridial spores persist in the dairy farm environment indefinitely.

2.8 Conclusion

The thermoduric nature of anaerobic spores means they are not eliminated by pasteurisation. Spores belonging to the genus *Clostridium* are of significant relevance to the dairy industry, as this genus contains known human pathogens as well as bacteria involved in the spoilage of milk products. The majority of the pathogenic and spoilage *Clostridium* spp. of relevance to the dairy industry fall within the metabolically distinct SRC group, and can be tested for based on this phenotype. Regulations for the limits of SRCs in milk vary between countries and milk producers, but in general there is a low tolerance for this group of clostridia in milk.

Unfortunately, the conventional assays which have been applied to test for the presence of this group of microorganisms are time consuming and labour intensive and are gradually being replaced by molecular type assays which rely on the detection of nucleic acids by using PCR. However, care must be taken with these methods to differentiate between live and dead cells.

Finally, farming practices are perhaps the single most important factor in controlling anaerobic spore numbers in BTM. The use of good quality silage, adherence to stringent shed/cubicle, parlour/milking equipment cleaning routines and maintenance, as well as a rigorous udder cleaning and teat preparation prior to milking are all considered to be GFMP. The strict application of these GFMPs is necessary to reduce the risk of anaerobic sporeformers contaminating BTM.

2.9 REFERENCES

Allison, C. and G. T. Macfarlane (1990). "Regulation of protease production in Clostridium sporogenes." Applied and environmental microbiology **56**(11): 3485-3490.

Andersson, A., U. Rönner and P. E. Granum (1995). "What problems does the food industry have with the spore-forming pathogens Bacillus cereus and Clostridium perfringens?" International journal of food microbiology **28**(2): 145-155.

Anon (1982). "Disease attributed to dairy products." British Medical Journal **285(6355):1664**.

Arias, C., B. Oliete, S. Seseña, L. Jimenez, M. Pérez-Guzmán and R. Arias (2013). "Importance of on-farm management practices on lactate-fermenting Clostridium spp. spore contamination of Manchega ewe milk: Determination of risk factors and characterization of Clostridium population." Small Ruminant Research **111**(1): 120-128.

Arnon, S. S. (1980). "Infant botulism." Annual review of medicine **31**(1): 541-560.

Arnon, S. S. (1998). "Infant botulism." Textbook of Pediatric Infectious Diseases. 4th ed. Philadelphia, Pa: WB Saunders Co: 1570-1577.

Arnon, S. S., T. F. Midura, K. Damus, B. Thompson, R. M. Wood and J. Chin (1979). "Honey and other environmental risk factors for infant botulism." The Journal of Pediatrics **94**(2): 331-336.

Aureli, P. and G. Franciosa (2002). "Clostridium spp." Encyclopedia of Dairy Sciences **1**: 456-463.

Aureli, P., G. Franciosa and M. Pourshaban (1996). "Foodborne botulism in Italy." The Lancet **348**(9041): 1594.

Balslev, T., E. Østergaard, I. Madsen and D. Wandall (1997). "Infant botulism. The first culture-confirmed Danish case." Neuropediatrics **28**(05): 287-289.

Barash, J. R. and S. S. Arnon (2004). "Dual toxin-producing strain of Clostridium botulinum type Bf isolated from a California patient with infant botulism." Journal of clinical microbiology **42**(4): 1713-1715.

Barash, J. R., J. K. Hsia and S. S. Arnon (2010). "Presence of soil-dwelling clostridia in commercial powdered infant formulas." The Journal of pediatrics **156**(3): 402-408.

Barnes, C., J. T. Gerstle, M. H. Freedman and M. D. Carcao (2004). "Clostridium septicum myonecrosis in congenital neutropenia." Pediatrics **114**(6): e757-e760.

Bergère, J., P. Gouet, J. Hermier and G. Mocquot (1968). "Les Clostridium du groupe butyrique dans les produits laitiers." Ann. Inst. Pasteur Lille **19**: 41-54.

Billon, J., J. Guerin and M. Sebald (1980). "Toxinogenesis of Clostridium botulinum type B during maturation of soft cheese." Lait **60**(597): 329-342.

Blocher, J. and F. Busta (1983). "Bacterial spore resistance to acid." Food Technology **37**.

Brett, M., J. McLauchlin, A. Harris, S. O'Brien, N. Black, R. Forsyth, D. Roberts and F. Bolton (2005). "A case of infant botulism with a possible link to infant formula milk powder: evidence for the presence of more than one strain of Clostridium botulinum in clinical specimens and food." Journal of medical microbiology **54**(8): 769-776.

Brook, I. (1995). "Clostridial infection in children." Journal of medical microbiology **42**(2): 78-82.

Brook, I. (2007). "Infant botulism." Journal of Perinatology **27**(3): 175-180.

Brynestad, S. and P. E. Granum (2002). "Clostridium perfringens and foodborne infections." International journal of food microbiology **74**(3): 195-202.

Burgess, S. A., D. Lindsay and S. H. Flint (2010). "Thermophilic bacilli and their importance in dairy processing." International journal of food microbiology **144**(2): 215-225.

Campbell, L. L. and J. R. Postgate (1965). "Classification of the spore-forming sulfate-reducing bacteria." Bacteriological reviews **29**(3): 359.

Cano, R. J. and M. K. Borucki (1995). "Revival and identification of bacterial spores in 25-to 40-million-year-old Dominican amber." Science **268**(5213): 1060-1064.

Carlin, F., V. Broussolle, S. Perelle, S. Litman and P. Fach (2004). "Prevalence of Clostridium botulinum in food raw materials used in REPFEDs manufactured in France." International journal of food microbiology **91**(2): 141-145.

Cato, E., W. L. George and S. Finegold (1986). "Genus clostridium." Bergey's manual of systematic bacteriology **2**: 1141-1200.

Cavalcanti, M. T. H., T. Porto, A. L. F. Porto, I. V. Brandi, J. L. d. Lima Filho and A. Pessoa Junior (2004). "Large scale purification of Clostridium perfringens toxins: a review." Revista Brasileira de Ciências Farmacêuticas **40**(2): 151-164.

Charlebois, A., M. Jacques and M. Archambault (2014). "Biofilm formation of Clostridium perfringens and its exposure to low-dose antimicrobials." Frontiers in microbiology **5**.

Chen, L., R. M. Daniel and T. Coolbear (2003). "Detection and impact of protease and lipase activities in milk and milk powders." International dairy journal **13**(4): 255-275.

Christiansson, A., J. Bertilsson and B. Svensson (1999). " Bacillus cereus Spores in Raw Milk: Factors Affecting the Contamination of Milk During the Grazing Period." Journal of dairy science **82**(2): 305-314.

Cocolin, L., N. Innocente, M. Biasutti and G. Comi (2004). "The late blowing in cheese: a new molecular approach based on PCR and DGGE to study the microbial ecology of the alteration process." International Journal of Food Microbiology **90**(1): 83-91.

Cogan, T. (1977). "A review of heat resistant lipases and proteinases and the quality of dairy products." Irish Journal of Food Science and Technology: 95-105.

Collins-Thompson, D. L. and D. S. Wood (1992). "Control in dairy products." Clostridium botulinum. Ecology and Control in Food, New York: Marcel Dekker: 261-277.

Collins, M., P. Lawson, A. Willems, J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, H. Hippe and J. Farrow (1994). "The phylogeny of the genus Clostridium: proposal of five new genera and eleven new species combinations." International journal of systematic bacteriology **44**(4): 812-826.

Council, U. D. E. (2013). US Dairy Export Council.

Cousin, M. (1982). "Presence and activity of psychrotrophic microorganisms in milk and dairy products: a review." Journal of Food Protection (USA).

Cox, W. (1975). "Subject: Bitty cream and related problems: Problems associated with bacterial spores in heat-treated milk and dairy products." International Journal of Dairy Technology **28**(2): 59-68.

Cremonesi, P., L. Vanoni, T. Silvetti, S. Morandi and M. Brasca (2012). "Identification of Clostridium beijerinckii, Cl. butyricum, Cl. sporogenes, Cl. tyrobutyricum isolated from silage, raw milk and hard cheese by a multiplex PCR assay." Journal of Dairy Research **79**(03): 318-323.

Crielly, E., N. Logan and A. Anderton (1994). "Studies on the Bacillus flora of milk and milk products." Journal of applied bacteriology **77**(3): 256-263.

Critchley, E., P. Hayes and P. Isaacs (1989). "Outbreak of botulism in north west England and Wales, June, 1989." The Lancet **334**(8667): 849-853.

Derman, Y., M. Lindström, K. Selby and H. Korkeala (2011). "Growth of group II Clostridium botulinum strains at extreme temperatures." Journal of Food Protection® **74**(11): 1797-1804.

Dhaked, R. K., M. K. Singh, P. Singh and P. Gupta (2010). "Botulinum toxin: Bioweapon & magic drug." The Indian journal of medical research **132**(5): 489.

Dodds, K. L. (1993). "Clostridium botulinum in foods." Clostridium botulinum: ecology and control in foods: 53-68.

Dover, N., J. R. Barash, K. K. Hill, G. Xie and S. S. Arnon (2014). "Molecular characterization of a novel botulinum neurotoxin type H gene." Journal of Infectious Diseases **209**(2): 192-202.

Doyle, M. E. and K. Glass (2013). "Spores of Clostridium botulinum in Dried Dairy Products."

Durre, P. (2001). "From Pandora's box to cornucopia: clostridia—a historical perspective." Clostridia—biotechnology and medical applications. Wiley-VCH, New York, NY: 1-17.

Eklund, M., D. Wieler and F. Poysky (1967). "Outgrowth and toxin production of nonproteolytic type B Clostridium botulinum at 3.3 to 5.6 C." Journal of bacteriology **93**(4): 1461.

Elmoslemany, A., G. Keefe, I. Dohoo and B. Jayarao (2009). "Risk factors for bacteriological quality of bulk tank milk in Prince Edward Island dairy herds. Part 2: Bacteria count-specific risk factors." Journal of dairy science **92**(6): 2644-2652.

FAO/WHO. (2004). "Enterobacter sakazakii and other microorganisms in powdered infant formula: meeting

report." Microbiological Risk Assessment Series 6. , from ftp://ftp.fao.org/docrep/fao/007/y5502e/y5502e00.pdf.

Federation, R. (2008). Federal Law on Technical Regulations for Milk and Milk Products, Russian Federation. **No 88-FZ**.

Fenicia, L., L. Da Dalt, F. Anniballi, G. Franciosa, S. Zanconato and P. Aureli (2002). "A case of infant botulism due to neurotoxigenic Clostridium butyricum type E associated with Clostridium difficile colitis." European Journal of Clinical Microbiology and Infectious Diseases **21**(10): 736-738.

Fischer, M., S. Zhu and E. de Ree (2012). "Culture Media for the Detection and Enumeration of Clostridia in Food." Handbook of Culture Media for Food and Water Microbiology: 66.

Fisher, D. J., K. Miyamoto, B. Harrison, S. Akimoto, M. R. Sarker and B. A. McClane (2005). "Association of beta2 toxin production with Clostridium perfringens type A human gastrointestinal disease isolates carrying a plasmid enterotoxin gene." Mol Microbiol **56**(3): 747-762.

Flint, S., P. Bremer and J. Brooks (1997). "Biofilms in dairy manufacturing plant-description, current concerns and methods of control." Biofouling **11**(1): 81-97.

Franciosa, G., M. Pourshaban, M. Gianfranceschi, A. Gattuso, L. Fenicia, A. M. Ferrini, V. Mannoni, G. De Luca and P. Aureli (1999). "Clostridium botulinum spores and toxin in mascarpone cheese and other milk products." Journal of Food Protection® **62**(8): 867-871.

Fritze, D. (2004). "Taxonomy of the genus Bacillus and related genera: the aerobic endospore-forming bacteria." Phytopathology **94**(11): 1245-1248.

Fuchs, A.-R. and G. Bonde (1957). "The nutritional requirements of Clostridium perfringens." Journal of general microbiology **16**(2): 317-329.

Garde, S., R. Arias, P. Gaya and M. Nuñez (2011). "Occurrence of Clostridium spp. in ovine milk and Manchego cheese with late blowing defect: Identification and characterization of isolates." International dairy journal **21**(4): 272-278.

Garde, S., P. Gaya, R. Arias and M. Nuñez (2012). "Enhanced PFGE protocol to study the genomic diversity of Clostridium spp. isolated from Manchego cheeses with late blowing defect." Food Control **28**(2): 392-399.

Gibbs, B. and B. Freame (1965). "Methods for the recovery of clostridia from foods." Journal of Applied Microbiology **28**(1): 95-111.

Gibert, M., C. Jolivet-Renaud and M. R. Popoff (1997). "Beta2 toxin, a novel toxin produced by Clostridium perfringens." Gene **203**(1): 65-73.

Gimenez, D. F. and A. S. Ciccarelli (1978). "New strains of Clostridium botulinum subtype Af." Zentralbl Bakteriol Orig A **240**(2): 215-220.

Gkiourtzidis, K., J. Frey, E. Bourtzi-Hatzopoulou, N. Iliadis and K. Sarris (2001). "PCR detection and prevalence of α-, β-, β2-, ε-, ι-and enterotoxin genes in Clostridium perfringens isolated from lambs with clostridial dysentery." Veterinary microbiology **82**(1): 39-43.

Gleeson, D., A. O'Connell and K. Jordan (2013). "Review of potential sources and control of thermoduric bacteria in bulk-tank milk." Irish Journal of Agricultural and Food Research **52**(2): 217-227.

Goudkov, A. and M. E. Sharpe (1965). "Clostridia in dairying." Journal of Applied Microbiology **28**(1): 63-73.

Graham, A., D. Mason, F. Maxwell and M. Peck (1997). "Effect of pH and NaCl on growth from spores of non-proteolytic Clostridium botulinum at chill temperature." Letters in applied microbiology **24**(2): 95-100.

Griffiths, M., J. Phillips, I. West and D. Muir (1988). "The effect of extended low-temperature storage of raw milk on the quality of pasteurized and UHT milk." Food microbiology **5**(2): 75-87.

Gui, L., C. Subramony, J. Fratkin and M. D. Hughson (2002). "Fatal enteritis necroticans (pigbel) in a diabetic adult." Modern pathology **15**(1): 66-70.

Hall, J., L. McCroskey, B. Pincomb and C. Hatheway (1985). "Isolation of an organism resembling Clostridium barati which produces type F botulinal toxin from an infant with botulism." Journal of clinical microbiology **21**(4): 654-655.

Hanson, M., W. Wendorff and K. Houck (2005). "Effect of heat treatment of milk on activation of Bacillus spores." Journal of Food Protection® **68**(7): 1484-1486.

Hassan, K. A., L. D. Elbourne, S. G. Tetu, S. B. Melville, J. I. Rood and I. T. Paulsen (2014). "Genomic analyses of Clostridium perfringens isolates from five toxinotypes." Research in Microbiology.

Hatheway, C. L. (1990). "Toxigenic clostridia." Clin Microbiol Rev **3**(1): 66-98.

Hatheway, C. L. and L. M. McCroskey (1987). "Examination of feces and serum for diagnosis of infant botulism in 336 patients." J Clin Microbiol **25**(12): 2334-2338.

Hippe, H., Andreesen, J. R. & Gottschalk, G. (1992). The genus Clostridium – nonmedical. The Prokaryotes, 2nd edn,. **vol. 2,:** pp. 1800–1866.

Huck, J., M. Sonnen and K. Boor (2008). "Tracking heat-resistant, cold-thriving fluid milk spoilage bacteria from farm to packaged product." Journal of dairy science **91**(3): 1218-1228.

Hussong, R. and B. Hammer (1930). "Anaerobic Bacteria in Dairy Products: I. Numbers of Spores of Anaerobic Bacteria in Milk and Cream." Journal of bacteriology **19**(2): 89.

ICMSF (2014). Usefulness of testing for Clostridium botulinum in powdered infant

formula and dairy-based ingredients for infant formula. ICMSF.

Jabbari, A., F. Tekyei, M. Esmaeilizad and R. Pilehchian Langroudi (2012). "Occurrence of Beta2 toxigenic Clostridium perfringens isolates with different toxin types in Iran." Archives of Razi **67**(2): 133-137.

Jayarao, B., S. Pillai, A. Sawant, D. Wolfgang and N. Hegde (2004). "Guidelines for monitoring bulk tank milk somatic cell and bacterial counts." Journal of Dairy Science **87**(10): 3561-3573.

Johnson, E. A., W. H. Tepp, M. Bradshaw, R. J. Gilbert, P. E. Cook and E. D. G. McIntosh (2005). "Characterization of Clostridium botulinum strains associated with an infant botulism case in the United Kingdom." Journal of clinical microbiology **43**(6): 2602-2607.

Josephson, K., C. Gerba and I. Pepper (1993). "Polymerase chain reaction detection of nonviable bacterial pathogens." Applied and Environmental Microbiology **59**(10): 3513-3515.

Julien, M.-C., P. Dion, C. Lafreniere, H. Antoun and P. Drouin (2008). "Sources of clostridia in raw milk on farms." Applied and environmental microbiology **74**(20): 6348-6357.

Juneja, V. and H. Thippareddi (2004). "Inhibitory effects of organic acid salts on growth of Clostridium perfringens from spore inocula during chilling of marinated ground turkey breast." International journal of food microbiology **93**(2): 155-163.

Kakinuma, H., H. Maruyama, H. Takahashi, K. Yamakawa and S. Nakamura (1996). "The first case of type B infant botulism in Japan." Pediatrics International **38**(5): 541-543.

Klijn, N., F. Nieuwenhof, J. D. Hoolwerf, C. Van Der Waals and A. H. Weerkamp (1995). "Identification of Clostridium tyrobutyricum as the causative agent of late blowing in cheese by species-specific PCR amplification." Applied and Environmental Microbiology **61**(8): 2919-2924.

Kosikowski, F. and P. Fox (1968). "Low heat, hydrogen peroxide, and bactofugation treatments of milk to control coliforms in cheddar cheese." Journal of Dairy Science **51**(7): 1018-1022.

Kramer, J. M. and R. J. Gilbert (1989). "Bacillus cereus and other Bacillus species." Foodborne bacterial pathogens **19**: 21-70.

Lahti, P., A. Heikinheimo, T. Johansson and H. Korkeala (2008). "Clostridium perfringens type A strains carrying a plasmid-borne enterotoxin gene (genotype IS1151-cpe or IS1470-like-cpe) as a common cause of food poisoning." J Clin Microbiol **46**(1): 371-373.

Le Bourhis, A.-G., J. Doré, J.-P. Carlier, J.-F. Chamba, M.-R. Popoff and J.-L. Tholozan (2007). "Contribution of< i> C. beijerinckii</i> and< i> C. sporogenes</i> in association with< i> C. tyrobutyricum</i> to the butyric fermentation in Emmental type cheese." International journal of food microbiology **113**(2): 154-163.

Li, J., V. Adams, T. L. Bannam, K. Miyamoto, J. P. Garcia, F. A. Uzal, J. I. Rood and B. A. McClane (2013). "Toxin plasmids of Clostridium perfringens." Microbiology and Molecular Biology Reviews **77**(2): 208-233.

Li, J. and B. A. McClane (2006). "Comparative effects of osmotic, sodium nitrite-induced, and pH-induced stress on growth and survival of Clostridium perfringens type A isolates carrying chromosomal or plasmid-borne enterotoxin genes." Appl Environ Microbiol **72**(12): 7620-7625.

Li, J. and B. A. McClane (2006). "Further comparison of temperature effects on growth and survival of Clostridium perfringens type A isolates carrying a chromosomal or plasmid-borne enterotoxin gene." Appl Environ Microbiol **72**(7): 4561-4568.

Lindström, M., A. Heikinheimo, P. Lahti and H. Korkeala (2011). "Novel insights into the epidemiology of Clostridium perfringens type A food poisoning." Food microbiology **28**(2): 192-198.

Lindström, M., R. Keto, A. Markkula, M. Nevas, S. Hielm and H. Korkeala (2001). "Multiplex PCR assay for detection and identification of Clostridium botulinum types A, B, E, and F in food and fecal material." Applied and environmental microbiology **67**(12): 5694-5699.

López-Enríquez, L., D. Rodríguez-Lázaro and M. Hernández (2007). "Quantitative detection of Clostridium tyrobutyricum in milk by real-time PCR." Applied and environmental microbiology **73**(11): 3747-3751.

Ludwig, W., K.-H. Schleifer and W. B. Whitman (2009). Revised road map to the phylumFirmicutes. Bergey's Manual® of Systematic Bacteriology, Springer**:** 1-13.

Magnusson, M., A. Christiansson and B. Svensson (2007). "Bacillus cereus Spores During Housing of Dairy Cows: Factors Affecting Contamination of Raw Milk." Journal of dairy science **90**(6): 2745-2754.

Magnusson, M., A. Christiansson, B. Svensson and C. Kolstrup (2006). "Effect of different premilking manual teat-cleaning methods on bacterial spores in milk." Journal of dairy science **89**(10): 3866-3875.

Mahant, N., P. Clouston and I. Lorentz (2000). "The current use of botulinum toxin." Journal of clinical neuroscience **7**(5): 389-394.

Manafi, M. (2000). "New developments in chromogenic and fluorogenic culture media." International Journal of Food Microbiology **60**(2): 205-218.

Martin, S. (2003). "Clostridium botulinumtype D intoxication in a dairy herd in Ontario." The Canadian Veterinary Journal **44**(6): 493.

May, M., M. Coulthard, G. Delbridge, J. McEniery, C. Burke and M. Nissen (2002). "Difficulties in the diagnosis and management of infant botulism." J Paediatr Child Health **38**(4): 425-426; author reply 426-427.

McAuley, C. M., K. McMillan, S. C. Moore, N. Fegan and E. M. Fox (2014). "Prevalence and characterization of foodborne pathogens from Australian dairy farm environments." Journal of dairy science.

McDonnell, A. and L. Holmes (1990). "Haemoglobinuria due to Clostridium perfringens type A mastitis in a ewe." British Veterinary Journal **146**(4): 380-381.

McKinnon, C. H. and G. L. Pettipher (1983). "A survey of sources of heat-resistant bacteria in milk with particular reference to psychrotrophic spore-forming bacteria." Journal of dairy research **50**(02): 163-170.

Mead, G. (1969). "The Use of Sulphite-containing Media in the Isolation of Clostridium welchii." Journal of Applied Microbiology **32**(3): 358-361.

Medini, D., C. Donati, H. Tettelin, V. Masignani and R. Rappuoli (2005). "The microbial pan-genome." Current opinion in genetics & development **15**(6): 589-594.

Melin, M., H. Wiktorsson, A. Christiansson, J. McLean, M. Sinclair and B. West (2002). Teat cleaning efficiency before milking in DeLaval VMSTM versus conventional manual cleaning, using Clostridium tyrobutyricum spores as marker. First North American Conference on robotic milking, Toronto, Canada, 20-22 March, 2002., Wageningen Pers.

Meyer, K. and B. Eddie (1951). "Perspectives concerning botulism." Zeitschrift für Hygiene und Infektionskrankheiten **133**(4): 255-263.

Minamisawa, K., K. Nishioka, T. Miyaki, B. Ye, T. Miyamoto, M. You, A. Saito, M. Saito, W. L. Barraquio and N. Teaumroong (2004). "Anaerobic nitrogen-fixing consortia consisting of clostridia isolated from gramineous plants." Applied and environmental microbiology **70**(5): 3096-3102.

Moschonas, G., D. Bolton, J. Sheridan and D. McDowell (2010). "The effect of storage temperature and inoculum level on the time of onset of 'blown pack'spoilage." Journal of applied microbiology **108**(2): 532-539.

Murphy, P. M., D. Lynch and P. M. Kelly (1999). "Growth of thermophilic spore forming bacilli in milk during the manufacture of low heat powders." International journal of dairy technology **52**(2): 45-50.

Murphy, S. and K. Boor (2000). "Trouble-shooting sources and causes of high bacteria counts in raw milk." Dairy, Food and Environmental Sanitation **20**(8): 606-611.

Nakamura, K., T. Kohda, K. Umeda, H. Yamamoto, M. Mukamoto and S. Kozaki (2010). "Characterization of the D/C mosaic neurotoxin produced by Clostridium botulinum associated with bovine botulism in Japan." Veterinary microbiology **140**(1): 147-154.

Nocker, A., P. Sossa-Fernandez, M. D. Burr and A. K. Camper (2007). "Use of propidium monoazide for live/dead distinction in microbial ecology." Applied and environmental microbiology **73**(16): 5111-5117.

O'Mahony, M., E. Mitchell, R. Gilbert, D. Hutchinson, N. Begg, J. Rodhouse and J. Morris (1990). "An outbreak of foodborne botulism associated with contaminated hazelnut yoghurt." Epidemiology and infection **104**(03): 389-395.

Oh, M.-H., J.-S. Ham and J. M. Cox (2012). "Diversity and toxigenicity among members of the Bacillus cereus group." International journal of food microbiology **152**(1): 1-8.

Osman, K., M. El-Enbaawy, N. Ezzeldeen and H. Hussein (2009). "Mastitis in dairy buffalo and cattle in Egypt due to Clostridium perfringens: prevalence, incidence, risk factors and costs." Revue scientifique et technique (International Office of Epizootics) **28**(3): 975-986.

Pahlow, G., R. Muck, F. Driehuis, S. Oude Elferink, S. spoelstra, D. Buxton and M. RE (2003). "Silage science and technology." Silage science and technology.

Pankey, J. (1989). "Premilking udder hygiene." Journal of dairy science **72**(5): 1308-1312.

Peck, M. W., J. Plowman, C. F. Aldus, G. M. Wyatt, W. P. Izurieta, S. C. Stringer and G. C. Barker (2010). "Development and application of a new method for specific and sensitive enumeration of

spores of nonproteolytic Clostridium botulinum types B, E, and F in foods and food materials." Appl Environ Microbiol **76**(19): 6607-6614.

Phillips, J. and M. Griffiths (1986). "Factors contributing to the seasonal variation of Bacillus spp. in pasteurized dairy products." Journal of applied bacteriology **61**(4): 275-285.

Pickett, J., B. Berg, E. Chaplin and M.-A. Brunstetter-Shafer (1976). "Syndrome of botulism in infancy: clinical and electrophysiologic study." The New England journal of medicine **295**(14): 770-772.

Prevost, S., J.-L. Cayol, F. Zuber, J.-L. Tholozan and F. Remize (2013). "Characterization of clostridial species and sulfite-reducing anaerobes isolated from foie gras with respect to microbial quality and safety." Food Control **32**(1): 222-227.

Prevot, A. (1948). Recherches sur la réduction des sulfates et des sulfites minéraux par les bactéries anaérobies. Annales de L Institut Pasteur, Masson Editeur 120 Blvd Saint-Germain, 75280 Paris 06, FRANCE.

Prévot, V., F. Tweepenninckx, E. Van Nerom, A. Linden and A. Kimpe (2007). "Optimization of polymerase chain reaction for detection of Clostridium botulinum type C and D in bovine samples." Zoonoses and public health **54**(8): 320-327.

Princewell, T. and M. Agba (1982). "Examination of bovine faeces for the isolation and identification of Clostridium species." Journal of Applied Microbiology **52**(1): 97-102.

Rammer, C. (1996). "Quality of grass silage infected with spores of Clostridium tyrobutyricum." Grass and Forage Science **51**(1): 88-95.

Ranieri, M., J. Huck, M. Sonnen, D. Barbano and K. Boor (2009). "High temperature, short time pasteurization temperatures inversely affect bacterial numbers during refrigerated storage of pasteurized fluid milk." Journal of dairy science **92**(10): 4823-4832.

Ranieri, M. L., R. A. Ivy, W. R. Mitchell, E. Call, S. N. Masiello, M. Wiedmann and K. J. Boor (2012). "Real-time PCR detection of Paenibacillus spp. in raw milk to predict shelf life performance of pasteurized fluid milk products." Applied and environmental microbiology **78**(16): 5855-5863.

Reinemann, D. J., G. Wolters, P. Billon, O. Lind and M. D. Rasmussen (2003). "Review of practices for cleaning and sanitation of milking machines." Bulletin-International Dairy Federation: 3-18.

Ribeiro, M., G. Lara, S. Bicudo, A. Souza, T. Salerno, A. Siqueira and J. Geraldo (2007). "An unusual gangrenous goat mastitis caused by Staphylococcus aureus, Clostridium perfringens and Escherichia coli co-infection." Arquivo Brasileiro de Medicina Veterinária e Zootecnia **59**(3): 810-812.

Roberts, T. and A. Hitchins (1969). "Resistance of spores." The bacterial spore **2**.

Robinson, A. and P. Manser (1977). "Mastitis in a heifer caused by Clostridium welchii, type" A"." Veterinary Record **101**(2): 37-37.

Rogers, P. and G. Gottschalk (1993). "Biochemistry and regulation of acid and solvent production in clostridia." BIOTECHNOLOGY SERIES: 25-25.

Ronimus, R. S., L. E. Parker, N. Turner, S. Poudel, A. Rückert and H. W. Morgan (2003). "A RAPD-based comparison of thermophilic bacilli from milk powders." International journal of food microbiology **85**(1): 45-61.

Ronimus, R. S., A. Rueckert and H. W. Morgan (2006). "Survival of thermophilic spore-forming bacteria in a 90+ year old milk powder from Ernest Shackelton's Cape Royds Hut in Antarctica." J Dairy Res **73**(2): 235-243.

Russell, A. (1990). "Bacterial spores and chemical sporicidal agents." Clinical Microbiology Reviews **3**(2): 99-119.

Rysstad, G. and J. Kolstad (2006). "Extended shelf life milk—advances in technology." International journal of dairy technology **59**(2): 85-96.

Santos-Buelga, J. A., M. D. Collins and A. K. East (1998). "Characterization of the genes encoding the botulinum neurotoxin complex in a strain of Clostridium botulinum producing type B and F neurotoxins." Current microbiology **37**(5): 312-318.

Sarker, M. R., R. P. Shivers, S. G. Sparks, V. K. Juneja and B. A. McClane (2000). "Comparative experiments to examine the effects of heating on vegetative cells and spores of Clostridium

perfringens isolates carrying plasmid genes versus chromosomal enterotoxin genes." Applied and environmental microbiology **66**(8): 3234-3240.

Savoy, d. G. G., d. V. G. Font, d. R. H. A. Pesce and G. Oliver (1981). "Isolation and identification of anaerobic contaminants from a machine for producing processed cheese." Revista Argentina de microbiologia **14**(2): 105-110.

Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M.-A. Widdowson, S. L. Roy, J. L. Jones and P. M. Griffin (2011). "Foodborne illness acquired in the United States—major pathogens." Emerg Infect Dis **17**(1).

Schmidt-Dannert, C., M. L. Rua and R. D. Schmid (1997). "Two novel lipases from thermophile Bacillus thermocatenulatus: screening, purification, cloning, overexpression, and properties." Methods Enzymol **284**: 194-220.

Schreiner, D. and P. Ruegg (2003). "Relationship between udder and leg hygiene scores and subclinical mastitis." Journal of dairy science **86**(11): 3460-3465.

Sebald, M., J. Jouglard and G. Gilles (1974). B botulism in man due to cheese (author's transl). Annales de microbiologie.

Shapiro, R. L., C. Hatheway and D. L. Swerdlow (1998). "Botulism in the United States: a clinical and epidemiologic review." Annals of Internal Medicine **129**(3): 221-228.

Silva, A. R., E. N. Paulo, A. S. Sant'Ana, R. D. Chaves and P. R. Massaguer (2011). "Involvement of Clostridium gasigenes and C. algidicarnis in 'blown pack'spoilage of Brazilian vacuum-packed beef." International journal of food microbiology **148**(3): 156-163.

Skanderby, M., V. Westergaard, A. Partridge and D. Muir (2009). "5 Dried Milk Products." Dairy Powders and Concentrated Products: 180.

SkjelkvÅLe, R. and T. Uemura (1977). "Experimental Diarrhoea in Human Volunteers Following Oral Administration of Clostridium perfringens Enterotoxin." Journal of Applied Bacteriology **43**(2): 281-286.

Slaghuis, B. A., M. C. Te Giffel, R. R. Beumer and G. André (1997). "Effect of pasturing on the incidence of Bacillus cereus spores in raw milk." International dairy journal **7**(4): 201-205.

Smokvina, T., M. Wels, J. Polka, C. Chervaux, S. Brisse, J. Boekhorst, J. E. van Hylckama Vlieg and R. J. Siezen (2013). "Lactobacillus paracasei Comparative Genomics: Towards Species Pan-Genome Definition and Exploitation of Diversity." PloS one **8**(7): e68731.

Sobel, J. (2005). "Botulism." Clinical Infectious Diseases **41**(8): 1167-1173.

Song, Y., C. Liu and S. M. Finegold (2004). "Real-time PCR quantitation of clostridia in feces of autistic children." Applied and Environmental Microbiology **70**(11): 6459-6465.

Sørhaug, T. and L. Stepaniak (1997). "Psychrotrophs and their enzymes in milk and dairy products: quality aspects." Trends in Food Science & Technology **8**(2): 35-41.

Stadhouders, J. and K. Jørgensen (1990). "Prevention of the contamination of raw milk by a hygienic milk production." Bulletin of the international dairy federation(251): 32-36.

Standards, I. O. o. (2003). Microbiology of food and animal feeding stuffs - Horizontal method for the enumeration of sulfite-reducing bacteria growing under anaerobic conditions. ISO 2003.

Standards, I. O. o. (2004). Microbiology of food and animal feeding stuffs -- Horizontal method for the enumeration of Clostridium perfringens -- Colony-count technique. ISO 2004.

Stiles, B. G., G. Barth, H. Barth and M. R. Popoff (2013). "Clostridium perfringens Epsilon Toxin: A Malevolent Molecule for Animals and Man?" Toxins **5**(11): 2138-2160.

Su, Y.-C. and S. C. Ingham (2000). "Influence of milk centrifugation, brining and ripening conditions in preventing gas formation by Clostridium spp. in Gouda cheese." International journal of food microbiology **54**(3): 147-154.

Suen, J. C., C. L. Hatheway, A. G. Steigerwalt and D. Brenner (1988). "Genetic confirmation of identities of neurotoxigenic Clostridium baratii and Clostridium butyricum implicated as agents of infant botulism." Journal of clinical microbiology **26**(10): 2191-2192.

Sugiyama, H. (1951). "Studies on factors affecting the heat resistance of spores of Clostridium botulinum." Journal of bacteriology **62**(1): 81.

Te Giffel, M., R. Beumer, P. Granum and F. Rombouts (1997). "Isolation and characterisation of Bacillus cereus from pasteurised milk in household refrigerators in the Netherlands." International journal of food microbiology **34**(3): 307-318.

Te Giffel, M., R. Beumer, B. Slaghuis and F. Rombouts (1995). "Occurrence and characterization of (psychrotrophic) Bacillus cereus on farms in the Netherlands." Nederlands melk en Zuiveltijdschrift **49**(2-3): 125-138.

Te Giffel, M., A. Wagendorp, A. Herrewegh and F. Driehuis (2002). "Bacterial spores in silage and raw milk." Antonie van Leeuwenhoek **81**(1-4): 625-630.

Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones and A. S. Durkin (2005). "Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial “pan-genome”." Proceedings of the National Academy of Sciences of the United States of America **102**(39): 13950-13955.

Ting, P. T. and A. Freiman (2004). "The story of Clostridium botulinum: from food poisoning to Botox." Clinical Medicine **4**(3): 258-261.

Uzal, F. A., J. C. Freedman, A. Shrestha, J. R. Theoret, J. Garcia, M. M. Awad, V. Adams, R. J. Moore, J. I. Rood and B. A. McClane (2014). "Towards an understanding of the role of Clostridium perfringens toxins in human and animal disease." Future microbiology **9**(3): 361-377.

Velthuis, A. and M. van Asseldonk (2010). Association between product quality control and process quality control of bulk milk. 9th Wageningen International Conference on Chain and Network Management. Wageningen, The Netherlands.

Vissers, M., F. Driehuis, M. Te Giffel, P. De Jong and J. Lankveld (2006). "Improving farm management by modeling the contamination of farm tank milk with butyric acid bacteria." Journal of dairy science **89**(3): 850-858.

Vissers, M., F. Driehuis, M. Te Giffel, P. De Jong and J. Lankveld (2007). "Concentrations of butyric acid bacteria spores in silage and relationships with aerobic deterioration." Journal of dairy science **90**(2): 928-936.

Vissers, M., F. Driehuis, M. Te Giffel, P. De Jong and J. Lankveld (2007). "Minimizing the level of butyric acid bacteria spores in farm tank milk." Journal of dairy science **90**(7): 3278-3285.

Vissers, M., M. Te Giffel, F. Driehuis, P. De Jong and J. Lankveld (2007). "Minimizing the Level of Bacillus cereus Spores in Farm Tank Milk." Journal of dairy science **90**(7): 3286-3293.

Waes, G. (1976). "Aerobic mesophilic spores in raw milk." Milchwissenschaft (Germany, FR)(31): 521-525.

Walstra, P., P. Walstra, J. T. Wouters and T. J. Geurts (2010). Dairy science and technology, CRC press.

Weenk, G., E. Fitzmaurice and D. Mossel (1991). "Selective enumeration of spores of Clostridium species in dried foods." Journal of Applied Microbiology **70**(2): 135-143.

Weenk, G., J. Van den Brink, C. Struijk and D. Mossel (1995). "Modified methods for the enumeration of spores of mesophilic Clostridium species in dried foods." International journal of food microbiology **27**(2): 185-200.

Wiegel, J., R. Tanner and F. A. Rainey (2006). An introduction to the family Clostridiaceae. The prokaryotes, Springer**:** 654-678.

Wilson, W. J. and E. M. M. V. Blair (1924). "The application of a sulphite-glucose-irobn agar medium to the quantitative estimation of B. welchii and other reducing bacteria in water supplies." The Journal of Pathology and Bacteriology **27**(1): 119-121.

Xiao, Y., A. Wagendorp, T. Abee and M. H. Wells-Bennik (2014). "Differential outgrowth potential of Clostridium perfringens food borne isolates with various cpe-genotypes in vacuum-packed ground beef during storage at 12° C." International journal of food microbiology.

Xu, D. and J.-C. Côté (2003). "Phylogenetic relationships between Bacillus species and related genera inferred from comparison of 3′ end 16S rDNA and 5′ end 16S–23S ITS nucleotide sequences." International Journal of Systematic and Evolutionary Microbiology **53**(3): 695-704.

Yan, R.-T., C.-X. Zhu, C. Golemboski and J.-S. Chen (1988). "Expression of solvent-forming enzymes and onset of solvent production in batch cultures of Clostridium beijerinckii ("Clostridium butylicum")." Applied and environmental microbiology **54**(3): 642-648.

Yutin, N. and M. Y. Galperin (2013). "A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia." Environmental microbiology **15**(10): 2631-2641.

Chapter 3

High-throughput metataxonomic characterisation of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions

Included as published in the International Journal of Food Microbiology (doi: 10.1016/j.ijfoodmicro.2017.05.019)

3.0 Abstract

Low temperature is used to control the growth of bacteria in milk, both pre- and post-pasteurisation. As the duration of refrigerated storage extends, psychrotrophs dominate the milk microbiota, that can produce heat stable lipases which negatively impact the organoleptic qualities of milk. Here we examine the influence that refrigeration temperature (2 ^{o}C, 4 ^{o}C and 6^{o}C) and storage duration (96 h) have on the microbiota composition (16S profiling) of raw bulk tank milk (BTM). To reflect a proposed change to current farming practices, raw milk was blended after each milking (8 milkings) and stored for five consecutive days in each temperature-specific tank. Here 16S rRNA-based microbiota compositional analysis was performed after milk was collected on day 1 and again after the final addition of milk at day 5. In addition to assessing the impact of the duration and temperature of storage, the influence of lactation stage, i.e. mid- versus late-lactation, on the microbiota of the blended BTM was also examined. Overall, both temperature and length of storage had surprisingly little influence on the raw milk microbiota, other than an increase in proportions of *Gammaproteobacteria* in the blended milk samples collected after pooling on day 5, and in samples stored at 6^{o}C. However, lactation stage had a considerable influence on microbiota composition, with milk from mid-lactation containing higher proportions of *Bacteroides*, *Faecalibacterium*, *Campylobacter* and *Rhodanobacter,* and late-lactation milk containing higher proportions of Actinobacteria. Overall, the study demonstrates that current temperature and storage duration practises impact the microbiota of raw milk, but these impacts are modest relative to the more considerable differences between mid and late-lactation milk.

3.1 Introduction

The microbiota of raw milk is complex (Quigley, O'Sullivan et al. 2013), and its composition, which is influenced by a multitude of intrinsic and extrinsic factors, is an important consideration for milk producers, processors and consumers. Indeed, the microbiota of milk influences the subsequent production of a wide variety of dairy products, such as cheese, butter, yogurt and dairy powders, and can contribute to the quality and safety of these foods (McInnis, Kalanetra et al. 2015). Dairy producers therefore need to be aware of the influence of environmental factors, such as lactation period (McInnis, Kalanetra et al. 2015, O'Connell, Ruegg et al. 2016) and storage conditions, such as temperature and duration of storage, on the microbial composition of raw milk (O'Connell, Ruegg et al. 2016).

Currently, most of what is understood about the presence of undesirable microorganisms in milk has been elucidated from selective plate cultivation-based techniques. These culture-based assays reveal the presence or absence of specific groups of bacteria, based on their phenotype (Quigley, O'Sullivan et al. 2013). These phenotypic assays, which are most commonly utilised by the dairy industry, target bacteria that proliferate during cold storage (psychrotrophs) or survive heat treatments (thermoduric bacteria including spore-formers). Psychrotrophic populations, which may increase during storage at refrigeration temperatures, include *Pseudomonas* and *Acinetobacter* spp. (Raats, Offek et al. 2011, Quigley, O'Sullivan et al. 2013). These populations are of particular significance as they are primarily responsible for spoilage of refrigerated dairy products (Raats, Offek et al. 2011, Machado, Bazzolli et al. 2013), most frequently through the production of heat stable lipases which can survive heat treatments designed to eliminate psychrotrophic bacteria

(Andersson, Hedlund et al. 1979, Sørhaug and Stepaniak 1997). Thermoduric bacteria are also of concern due to their spoilage and toxigenic potential (Doyle, Gleeson et al. 2015).

Recently, it was found by culture-based surveillance that the microbial quality of blended raw milk stored at refrigeration temperatures (2, 4 or 6°C) was not significantly altered by storage time (O'Connell, Ruegg et al. 2016). However, a corresponding study that focused on lactation stage revealed that it has a more considerable influence, with total bacterial counts (TBCs) being higher in late lactation milk (O'Connell, McParland et al. 2015), which, in the Irish dairy farm system, corresponds to winter. These studies are of significant applied value because longer raw milk collection interval extensions are more practical for milk processors, storage at higher temperatures is more economic for milk producers, and reductions in the quality in late lactation milk can influence its downstream use. Despite the potential value of these findings, it is important to note that culture-based methods are ultimately limited to revealing what can be grown in laboratory conditions, which may represent only a fraction of the bacteria present in the environment (Ward, Bateson et al. 1992, Hugenholtz and Pace 1996). Advances in DNA-based technologies and, more specifically, the application of next generation sequencing has provided a greater insight into the microbiota composition of milk and dairy products (Ercolini, Russo et al. 2009, Verdier-Metz, Michel et al. 2009, Raats, Offek et al. 2011, Vacheyrou, Normand et al. 2011, Quigley, McCarthy et al. 2013). This type of molecular analysis was initially developed for environmental microbiology but is equally applicable to the analysis of raw milk and other dairy products (Thierry, Maillard et al. 2005, Mallet, Guéguen et al. 2012, Quigley, McCarthy et al. 2013, Wolfe, Button et al. 2014, McInnis, Kalanetra et al. 2015, Gschwendtner, Alatossava et al. 2016, Quigley, O'Sullivan et al. 2016, Walsh, Crispie et al. 2016)This present

study was run concurrently with O'Connell and colleagues (O'Connell, Ruegg et al. 2016). As such, the conditions and experimental design described here are identical as that study, with the exception of the way in which samples were processed for analysis and the goal of the study. The goal of this study was to characterise the raw milk microbiotausing high-throughput sequencing, while O'Connell and colleagues targeted a subset of cultivable microbes. Here, we address the important issues of storage duration, storage temperature and lactation period on the microbial content of raw milk using high-throughput metataxonomic analysis.

3.2 Methods

3.2.1 Experimental design

The study was conducted at the Animal and Grassland Research and Innovation Centre, Teagasc, Moorepark, Cork, Ireland, using milk produced from spring-calved dairy cows, as described previously by O'Connell and colleagues (Ercolini 2013).Milk production over two 6-week periods was studied; period 1 extended from August 11 to September 26, corresponding to mid-lactation, and from October 13 to November 21, corresponding to late-lactation. During period 1 and the first 4 weeks of period 2, the cows were outdoors consuming a diet of grass. During the remaining 2 weeks of period 2, the cows were housed indoors during times of heavy rainfall on cubicles fitted with rubber mats that were bedded with lime, and they consumed a diet consisting of approximately 50% grazed grass and 50% grass silage. Teats were disinfected prior to milking as described previously (O'Connell,

Ruegg et al. 2016). Two milking's were conducted daily for the duration this study. Upon the completion of each milking, equipment was sanitised as described previously (O'Connell, Ruegg et al. 2016). Three identical 4,000 L bulk tank units (Swiftcool, Dairymaster) were used in this study. The 3 bulk tanks were set to cool milk to the different temperatures at the beginning of each test period. Valves in the milk-line were used to divide the milk flow in equal proportions (300 L into each tank at each milking) to each of the 3 tanks. The milk passed through a plate cooler and was cooled to approximately 14.5°C before entering each tank. The milk was subsequently cooled to the desired temperature, 2, 4 or 6^{o}C, within the tank. Upon completion of the 96-h storage period, each bulk tank was sterilised as described previously (O'Connell, Ruegg et al. 2016)

Equal volumes of milk were pumped (300 L) into each tank at each milking for four days (n=8 milkings) each week, for two 6-week periods, representing mid and late-lactation milk, respectively, and each tank was set at a different temperature (2, 4 or 6^{o}C) at the beginning of each week. Each treatment was applied to each tank on two occasions within each period. Milk was collected aseptically from each bulk tank after the morning milking on day 1 and on day 5 (representing 96 h of storage) using sterile blue dippa collection bottles (OCON Chemicals, Ireland). The latter represented a mixture of all milk collected over the five day period and was investigated to assess the consequences of extending milk collection intervals at farms.

3.2.2 DNA extraction

For each sample 15 mL of raw milk was centrifuged at 5444 xg for 30 minutes at 4 ^{o}C. The fat layer was carefully removed and the supernatant was decanted. Cell pellets were then homogenised in 90 µL lysozyme solution 50mg/mL (Sigma Aldrich, Arklow, Co. Wicklow, Ireland) and 50 uL of 50U/mL mutanolysin (Sigma Aldrich, Arklow, Co. Wicklow, Ireland), vortexed and incubated at 55 ^{o}C for 15 minutes vortexed at 2-3min intervals. Then 28 µL of proteinase K solution (Sigma Aldrich, Arklow, Co. Wicklow, Ireland) was then added to the cell pellet homogenate and the samples were incubated at 55 ^{o}C for 15 minutes. After incubation samples were centrifuged at 14, 000 x g for 5 minutes, supernatant was removed and the PowerFood DNA isolation kit was used as per manual (Mobio, Carlsbad CA) (O'Sullivan, Fallico et al. 2015). DNA was quantified and quality checked by gel electrophoresis and Nanodrop 1000 instrument (Thermo Fisher Scientific, Inc.).

3.2.3 Quantitative PCR

Quantitative PCR (qPCR) was carried out on samples to quantify the total bacteria in each sample. This qPCR was carried out as per (Fouhy, Guinane et al. 2012), except for the use of Kapa SYBR fast. Standards, samples and negative controls were all run in triplicate.

3.2.4 16S rRNA amplicon preparation and high throughput sequencing

The V3-V4 variable region of the 16S rRNA gene was amplified from DNA extracts using the 16S metagenomic sequencing library protocol (Illumina). PCR reactions were completed on the template DNA. Initially, the DNA was amplified with primers specific to the V3-V4 region

of the 16S rRNA gene which also incorporates the Illumina overhang adaptor (Walsh, Crispie et al. 2016). Samples were sequenced on the MiSeq sequencing platform in the Teagasc sequencing facility, using a 2 x 300 cycle V3 kit, following standard Illumina sequencing protocols.

3.2.5 Bioinformatic and statistical analysis

Three hundred base pair paired-end reads were assembled using FLASH (FLASH: fast length adjustment of short reads to improve genome assemblies). Further processing of paired-end reads including quality filtering based on a quality score of > 25 and removal of mismatched barcodes and sequences below length thresholds was completed using QIIME (Caporaso, Kuczynski et al. 2010). Denoising, chimera detection and clustering into operational taxonomic units (OTUs) (97% identity) were performed using USEARCH v7 (64-bit) (Edgar 2010). OTUs were aligned using PyNAST (PyNAST: python nearest alignment space termination; a flexible tool for aligning sequences to a template alignment) and taxonomy was assigned using BLAST against the SILVA SSURef database release 111. Samples were then rarefied to an even depth of sequences per sample. Alpha and beta diversities were generated using Phyloseq (McMurdie and Holmes 2013) package in R. The Wilcox rank sum test was run in R with Phyloseq to compare significant difference between sample day (1 or 5) and for lactation period (mid and late). P values were corrected for false discovery using the Benjamini Hochberg (BH) method (Benjamini and Hochberg 1995).

3.3 Results

3.3.1 Sequences

After DNA extraction, 16S rRNA amplicon generation, sequencing and quality filtering, an average of 53,901 reads were generated per sample, resulting in a mean of 9481 OTUs per sample. Reads were deposited in the European Nucleotide Archive database under accession number PRJEB16770.

3.3.2 Lactation stage has the most considerable effect on the alpha and beta diversity of the raw milk microbiota

In the majority of cases, no significant difference in alpha diversity was observed between milk samples when they were grouped by temperature and duration of storage. In one exceptional instance, there was a significant difference observed between day 1 and 5 samples stored at 6^{o}C during mid-lactation, when assessed using the chao1 index (p=0.025) and the Phylogenetic Diversity whole tree index (P=0.025; [supplementary table 1]). In this instance, alpha diversity was reduced in day five samples. In contrast, clear differences were apparent when samples were grouped by lactation stage, in that mid-lactation samples had a significantly greater alpha diversity, regardless of the index used (chao1 [p=0.001], Simpson diversity index [p=0.033], Shannon index [p=0.033], Observed species [p=0.001], PD whole tree index [p<0.001]; [supplementary table 1]).

The beta diversity of the raw milk microbiota was also investigated and presented in the form of principal coordinate analysis (PCoA). Samples partially cluster according to lactation stage in the Bray-Curtis PCoA plot (Fig.1), specifically samples from the last four weeks of

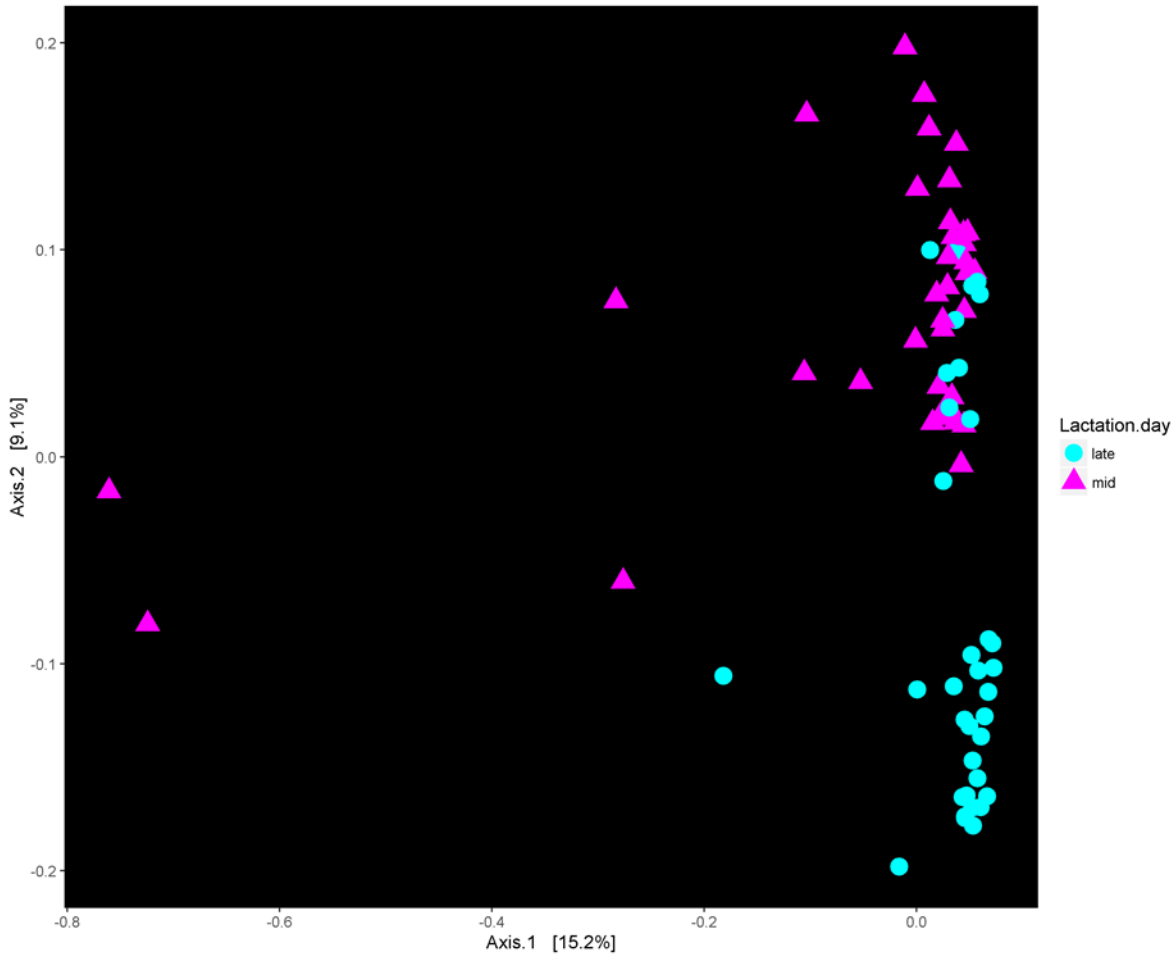


Fig.1: Bray-curtis distances of samples by lactation stage.

late-lactation cluster together (Fig S2). In contrast, no clustering on the basis of either temperature or sample day was observed (Supplementary Fig S1).

3.3.3 Taxonomic analysis highlights the influence of lactation stage and temperature on the raw milk microbiota

With the exception of *Ruminococcaceae* uncultured, the bacterial composition of raw milk is made up of taxa present at a mean relative abundance's lower than 10%. Furthermore, there is a preponderance of taxa present at a mean relative abundance of less than 0.5%. Additionally, it is noted that anaerobic taxa (*Clostridium*, *Clostridiales* Family XIII Incertae Sedis, *Lachnospiraceae* Incertae Sedis, *Ruminococcus* etc.) constitute the majority of the populations present above 1%.

Seventeen genera dominated the milk samples (i.e. were present at above 1%; Fig 2). Among taxa present above 0.05%, 85 taxa had significantly different relative abundances between mid and late-lactation (Fig S2). Of these, 21 taxa, including, for example, *Bacteroides*, *Faecalibacterium*, *Campylobacter* and *Rhodanobacter*, were significantly more abundant in mid-lactation. The remaining 64 were found to be more prevalent in late-lactation samples. These included some *Corynebacterium* spp. (*C. freneyi*, *C.*sp JY02 and *C.* sp MFC 5), *Micrococcus* sp. RNP02 and *Arthrobacter* sp. tsz11, all of which belong to the Phylum Actinobacteria (Fig S2). *Clostridium* was also found to be in higher abundance in late-lactation milk samples.

In contrast, no significant differences were found in proportions of taxonomic groups between day 1 and day 5 samples (where lactation stage and storage temperature was

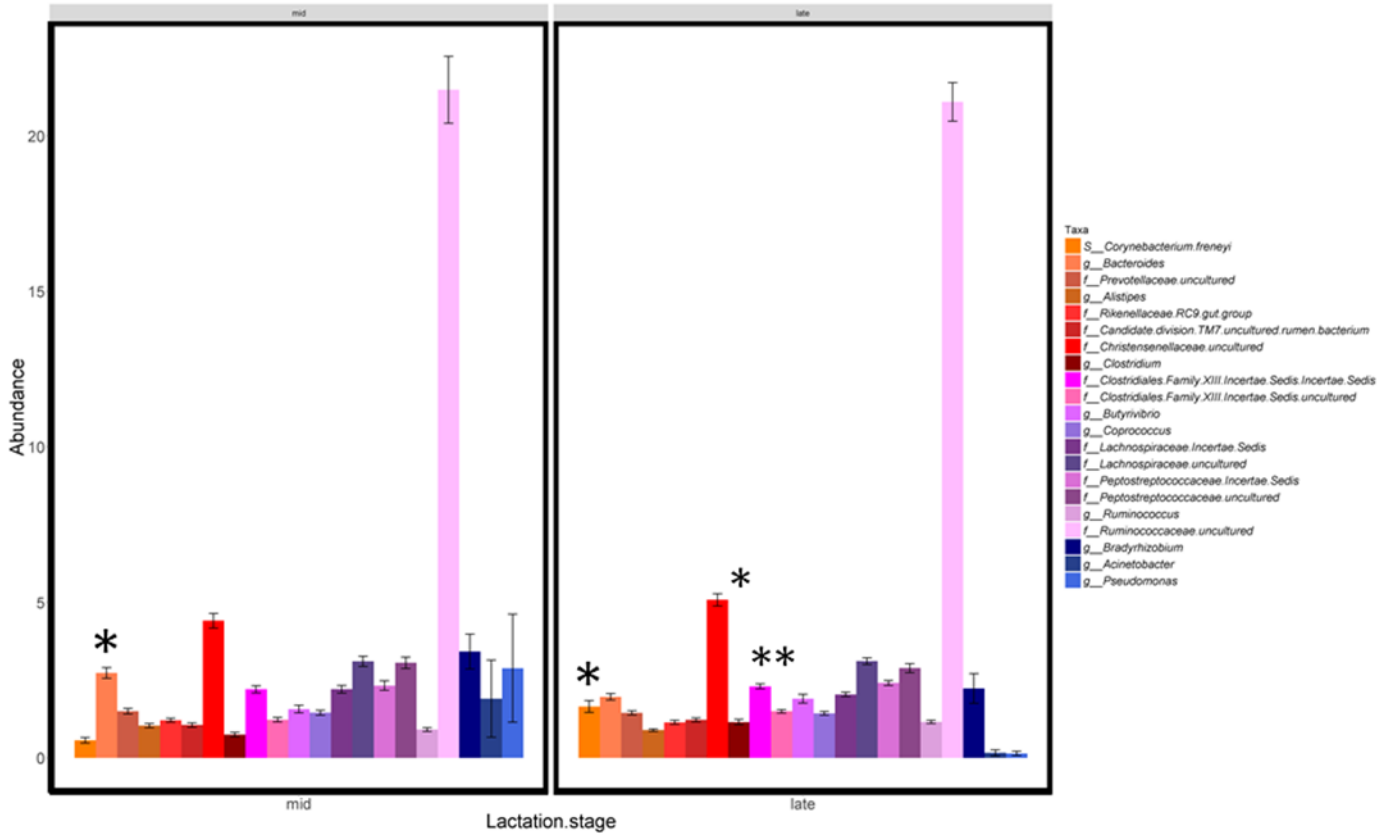


Fig.2: Taxa present above 1% in mid and late-lactation samples. Taxa which were significantly higher are denoted with an *. *Bacteroides* (P=<0.001), *Clostridium* (P=<0.001), *Corynebacterium freneyi* (P=<0.001), Family XIII *Incertae Sedis* uncultured (P=0.012), and *Ruminococcus* (P=0.005).

identical; Fig. S3-S8) or between samples stored at different temperatures (Fig. S9-S10). In 2^{o}C milk from mid-lactation, there were no observable differences in microbiota composition. However, for 4^{o}C milk from the same period there was an observed increase in proportions of *Streoptococcus* and *Pseudomonas* in day 5 samples relative to day 1 samples. A similar trend is seen in 6^{o}C samples from mid-lactation, where there is an observed increase in proportions in *Pseudomonas* and *Acinetobacter*, which was accompanied by a noticeable decrease in proportions of *Staphylococcus*, *Rhodanobacter* and *Ruminococcaceae* uncultured. This increase in *Pseudomonas* and *Acinetobacter* is also observed in late lactation samples stored at 6^{o}C, however the increase is much lower.

3.3.4 qPCR analysis to estimate total bacterial number in milk

The microbiota data presented above reflects the proportions of different taxa present. qPCR was employed to investigate total bacterial numbers, or more specifically, 16S rRNA gene copies. Total 16S rRNA copies were compared for BTM stored at 2, 4 and 6 ^{o}C between day 1 and day 5 (Table 1). During mid-lactation there were no significant differences in total bacterial numbers between day 1 and day 5 at any temperature. However, in late-lactation samples, a significant increase in total bacterial numbers occurred between day 1 and day 5 in BTM stored at 6 ^{o}C (P=0.011) (Table 1). These increases can be viewed in (Fig S12). Here, the increase in 16S copy numbers in 6 degrees at day 5 of mid and late-lactation can be observed relative to day 1. Higher 16S rRNA copy numbers were also observed in late-lactation milk relative to mid-lactation milk (Fig.S13).

Table 1: Total bacterial load† in milk as a function of storage temperature and lactation stage.

	Day 1	Day 5	P-value*
2°C mid-lactation	$7.95x10^4$	$1.15x10^5$	0.337
4°C mid-lactation	$8.42x10^4$	$1.31x10^5$	0.2
6°C mid-lactation	$1.15x10^5$	$5.06x10^5$	0.15
2°C late-lactation	$3.07x10^5$	$2.74x10^5$	1
4°C late-lactation	$3.56x10^5$	$5.68x10^5$	0.873
6°C late-lactation	$1.23x10^5$	$8.06x10^5$	0.011
	Mid	Late	P-value
Mid and late-lactation	111583	320500	<0.001

*P-value relates to differences between day 1 and day 5 values

†Determined by quantification of total copy number of 16S rRNA gene

3.4 Discussion

Here, high throughput sequencing was employed to assess the impact of seasonality, storage time and storage temperature on the microbial composition of blended raw milk. The objective of this study was to examine the impact of combining/blending milk collected over multiple days of milking on the raw milk microbiota. Thereby simulating a scenario in which BTM is collected by the dairy processor at less frequent intervals, while in parallel assessing the influence of storage temperature (2, 4 and 6^{o}C) and seasonality/lactation period (mid-/late-lactation).

One of the most striking observations from the study was the increase in proportions of OTUs (sequence-based bacterial classifications similar to traditional species) belonging to the Phylum Actinobacteria in samples from late-lactation relative to mid-lactation samples (Fig.S3). There was no screening conducted for Actinobacteria in the culture based analysis of O'Connell, therefore it is difficult to make comparisions. However this shift is in line with findings from a recent study which focused on the microbiota of raw milk from both wild type and genetically modified goats over a lactation cycle (McInnis, Kalanetra et al. 2015). This change may be due to environmental factors (temperature, humidity, weather, housing) or to physiological changes as a result of lactation stage. With respect to environmental factors, it is notable that animals are located in closer proximity to one another in late-lactation (when housed indoors), thus increasing the likelihood of transmission of skin and teat associated microbes, many of which are Actinobacteria, through the herd and subsequently into the raw milk. Species belonging to this phylum may contribute to the flavour development of dairy products by degrading proteins found in milk and cheese. However, while some species of the genus *Corynebacterium* can have a positive

influence on the maturation of cheese, it must be noted that other members of this genus are animal pathogens (Hogan, Smith et al. 1988, Fernandez-Garayzabal, Collins et al. 1997). Of the differences that were lactation period-dependent, the increase of Actinomycetales OTUs (*Corynebacterium* and *Micrococcus*) in late-lactation milk was particularly interesting, and is consistent with a recent investigation of the microbiota of goats milk (McInnis, Kalanetra et al. 2015) or milk from animals that have been housed indoors. The increase in *Clostridium* is important to consider as (O'Connell, Ruegg et al. 2016), also found a higher prevalence of SRCs in late-lactation milk samples in the culture based surveillance of these samples.

Of the bacteria that were found to be more abundant in mid-lactation, *Bacteroides* and *Faecalibacterium* are most frequently associated with the gastro-intestinal tract of mammals, but have also previously been found in culture-independent surveillance of raw milk (Quigley, McCarthy et al. 2013). *Rhodanobacter* is a member of the *Xanthomonadaceae*, which has previously been found in grass and soil samples from dairy farms (Doyle, Gleeson et al. 2017), possibly explaining the higher proportions of this bacteria in mid lactation. *Campylobacter* are recognised as pathogens that causes food-borne diarrhoea and have been previously been identified in raw buffalo milk using culture-independent approaches (Serraino, Florio et al. 2013). The source of this microbe is typically soil or water (Bronowski, James et al. 2014). The presence of *Campylobacter* is noteworthy as it was identified as being present in ~3% of raw milk samples in the Republic of Ireland recently (FSAI 2015). It is important to note that the nutrient content of milk produced by the herd differed by lactation stage (O'Connell, Ruegg et al. 2016), this change may have

had an influence on the microbial composition of the Raw milk and the total bacteria present.

In conclusion, the raw milk microbiota is dominated by bacteria present in low abundances this high-throughput metataxonomic sequencing. This highlights difference in the microbial compositions of milk during mid and late-lactation. There were considerably fewer differences between the microbiota of the samples with respect to storage temperature or storage duration. It is however important to consider that the raw milk being transferred to the bulk tanks was of considered to be of good microbiological quality (<4,800 cfu/mL) (O'Connell, Ruegg et al. 2016), due to the implementation of good hygiene practices. Storage temperature and duration can have more of a significant impact on the raw milk microbiota when milk is of poorer microbiological quality (e.g. 6^{o}C late-lactation). Results here indicate that lactation stage has a significant influence on milk microbiota composition possibly due to environmental exposure, and that BTM temperature and storage duration have a less apparent impact on the raw milk microbiota. This highlights the value of applying HTS based approaches to assess the impact that extrinsic factors have on the raw milk microbiota and provides insights that have the potential to benefit the agriculture and dairy processing industry sectors.

3.5 References

Andersson, R., C. Hedlund and U. Jonsson (1979). "Thermal inactivation of a heat-resistant lipase produced by the psychrotrophic bacterium Pseudomonas fluorescens." Journal of Dairy Science **62**(3): 361-367.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society. Series B (Methodological): 289-300.

Bronowski, C., C. E. James and C. Winstanley (2014). "Role of environmental survival in transmission of Campylobacter jejuni." FEMS Microbiology Letters **356**(1): 8-19.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich and J. I. Gordon (2010). "QIIME allows analysis of high-throughput community sequencing data." Nature methods **7**(5): 335-336.

Doyle, C. J., D. Gleeson, K. Jordan, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." International journal of food microbiology **197**: 77-87.

Doyle, C. J., D. Gleeson, P. W. O'Toole and P. D. Cotter (2017). "Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study." Applied and environmental microbiology **83**(2): e02694-02616.

Edgar, R. C. (2010). "Search and clustering orders of magnitude faster than BLAST." Bioinformatics **26**(19): 2460-2461.

Ercolini, D. (2013). "High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology." Applied and environmental microbiology **79**(10): 3148-3155.

Ercolini, D., F. Russo, I. Ferrocino and F. Villani (2009). "Molecular identification of mesophilic and psychrotrophic bacteria from raw cow's milk." Food Microbiology **26**(2): 228-231.

Fernandez-Garayzabal, J., M. Collins, R. Hutson, E. Fernandez, R. Monasterio, J. Marco and L. Dominguez (1997). "Corynebacterium mastitidis sp. nov., isolated from milk of sheep with subclinical mastitis." International journal of systematic bacteriology **47**(4): 1082-1085.

Fouhy, F., C. M. Guinane, S. Hussey, R. Wall, C. A. Ryan, E. M. Dempsey, B. Murphy, R. P. Ross, G. F. Fitzgerald, C. Stanton and P. D. Cotter (2012). "High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin." Antimicrobial Agents and Chemotherapy **56**(11): 5811-5820.

FSAI (2015). Raw Milk and Raw Milk Filter Microbiological Surveillance Programme, Food Safety Authority of Ireland.

Gschwendtner, S., T. Alatossava, S. Kublik, M. M. Fuka, M. Schloter and P. Munsch-Alatossava (2016). "N 2 Gas Flushing Alleviates the Loss of Bacterial Diversity and Inhibits Psychrotrophic Pseudomonas during the Cold Storage of Bovine Raw Milk." PloS one **11**(1): e0146015.

Hogan, J., K. Smith, D. Todhunter and P. Schoenberger (1988). "Rate of environmental mastitis in quarters infected with Corynebacterium bovis and Staphylococcus species." Journal of dairy science **71**(9): 2520-2525.

Hugenholtz, P. and N. R. Pace (1996). "Identifying microbial diversity in the natural environment: a molecular phylogenetic approach." Trends in biotechnology **14**(6): 190-197.

Machado, S. G., D. M. S. Bazzolli and M. C. D. Vanetti (2013). "Development of a PCR method for detecting proteolytic psychrotrophic bacteria in raw milk." International Dairy Journal **29**(1): 8-14.

Mallet, A., M. Guéguen, F. Kauffmann, C. Chesneau, A. Sesboué and N. Desmasures (2012). "Quantitative and qualitative microbial analysis of raw milk reveals substantial diversity influenced by herd management practices." International Dairy Journal **27**(1): 13-21.

McInnis, E. A., K. M. Kalanetra, D. A. Mills and E. A. Maga (2015). "Analysis of raw goat milk microbiota: Impact of stage of lactation and lysozyme on microbial diversity." Food microbiology **46**: 121-131.

McMurdie, P. J. and S. Holmes (2013). "phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data."
O'Connell, A., S. McParland, P. Ruegg, B. O'Brien and D. Gleeson (2015). "Seasonal trends in milk quality in Ireland between 2007 and 2011." Journal of Dairy Science **98**(6): 3778-3790.
O'Connell, A., P. Ruegg, K. Jordan, B. O'Brien and D. Gleeson (2016). "The effect of storage temperature and duration on the microbial quality of bulk tank milk." Journal of dairy science **99**(5): 3367-3374.
O'Sullivan, D. J., V. Fallico, O. O'Sullivan, P. L. McSweeney, J. J. Sheehan, P. D. Cotter and L. Giblin (2015). "High-throughput DNA sequencing to survey bacterial histidine and tyrosine decarboxylases in raw milk cheeses." BMC Microbiology **15**(1): 1.
Quigley, L., R. McCarthy, O. O'Sullivan, T. P. Beresford, G. F. Fitzgerald, R. P. Ross, C. Stanton and P. D. Cotter (2013). "The microbial content of raw and pasteurized cow's milk as determined by molecular approaches." Journal of Dairy Science.
Quigley, L., O. O'Sullivan, C. Stanton, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2013). "The complex microbiota of raw milk." FEMS Microbiology Reviews **37**(5): 664-698.
Quigley, L., D. J. O'Sullivan, D. Daly, O. O'Sullivan, Z. Burdikova, R. Vana, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. L. McSweeney (2016). "Thermus and the Pink Discoloration Defect in Cheese." mSystems **1**(3): e00023-00016.
Raats, D., M. Offek, D. Minz and M. Halpern (2011). "Molecular analysis of bacterial communities in raw cow milk and the impact of refrigeration on its structure and dynamics." Food microbiology **28**(3): 465-471.
Serraino, A., D. Florio, F. Giacometti, S. Piva, D. Mion and R. Zanoni (2013). "Presence of Campylobacter and Arcobacter species in in-line milk filters of farms authorized to produce and sell raw milk and of a water buffalo dairy farm in Italy." Journal of dairy science **96**(5): 2801-2807.
Sørhaug, T. and L. Stepaniak (1997). "Psychrotrophs and their enzymes in milk and dairy products: quality aspects." Trends in Food Science & Technology **8**(2): 35-41.
Thierry, A., M.-B. Maillard, R. Richoux, J.-R. Kerjean and S. Lortal (2005). "Propionibacterium freudenreichii strains quantitatively affect production of volatile compounds in Swiss cheese." Le Lait **85**(1-2): 57-74.
Vacheyrou, M., A.-C. Normand, P. Guyot, C. Cassagne, R. Piarroux and Y. Bouton (2011). "Cultivable microbial communities in raw cow milk and potential transfers from stables of sixteen French farms." International journal of food microbiology **146**(3): 253-262.
Verdier-Metz, I., V. Michel, C. Delbes and M.-C. Montel (2009). "Do milking practices influence the bacterial diversity of raw milk?" Food Microbiology **26**(3): 305-310.
Walsh, A. M., F. Crispie, K. Kilcawley, O. O'Sullivan, M. G. O'Sullivan, M. J. Claesson and P. D. Cotter (2016). "Microbial Succession and Flavor Production in the Fermented Dairy Beverage Kefir." mSystems **1**(5): e00052-00016.
Ward, D. M., M. M. Bateson, R. Weller and A. L. Ruff-Roberts (1992). Ribosomal RNA analysis of microorganisms as they occur in nature. Advances in microbial ecology, Springer**:** 219-286.
Wolfe, B. E., J. E. Button, M. Santarelli and R. J. Dutton (2014). "Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity." Cell **158**(2): 422-433.

Supplementary material

Table S1: Alpha diversity of all samples by group.

	Mid lactation	Late lactation	P-value
Chao1	1471	1314	0.001
Observed species	1062.5	972	0.001
PD whole tree	61.74	43.49	<0.001
Shannon	8.67	8.508	0.033
Simpson	0.9927	0.9917	0.033
	2°C day 1 mid lactation	**2°C day 5 mid lactation**	**P-value**
Chao1	1287	1378	0.337
Observed species	950.5	1042.5	0.2
PD whole tree	53.53	53.08	0.749
Shannon	8.488	8.766	0.2
Simpson	0.9921	0.9935	0.262
	4°C day 1 mid lactation	**4°C day 5 mid lactation**	**P-value**
Chao1	1071	1361	0.15
Observed species	837.5	998.5	0.078
PD whole tree	52.27	54.9	0.15
Shannon	8.332	8.633	0.15

Simpson	0.9912	0.9922	0.337

	6°C day 1 mid lactation	**6°C day 5 mid lactation**	**P-value**
Chao1	1413	1118	0.025
Observed species	982	887.5	0.078
PD whole tree	56.56	48.76	0.025
Shannon	8.436	8.264	0.423
Simpson	0.9887	0.9852	0.749

	2°C day 1 late lactation	**2°C day 5 late lactation**	**P-value**
Chao1	1651	1591	1
Observed species	1074	1072	0.936
PD whole tree	59.89	64.48	0.423
Shannon	8.543	8.682	0.337
Simpson	0.9916	0.9936	0.262

	4°C day 1 late lactation	**4°C day 5 late lactation**	**P-value**
Chao1	1432	1471	0.423
Observed species	1038	1068	0.521
PD whole tree	60.78	62.23	0.631
Shannon	8.56	8.628	0.262
Simpson	0.9919	0.9919	0.423

	6°C day 1 late lactation	6°C day 5 late lactation	P-value
Chao1	1466	1450	0.873
Observed species	1031	1066	1
PD whole tree	55.65	60.07	0.522
Shannon	8.804	8.0676	0.423
Simpson	0.9946	0.9926	0.423

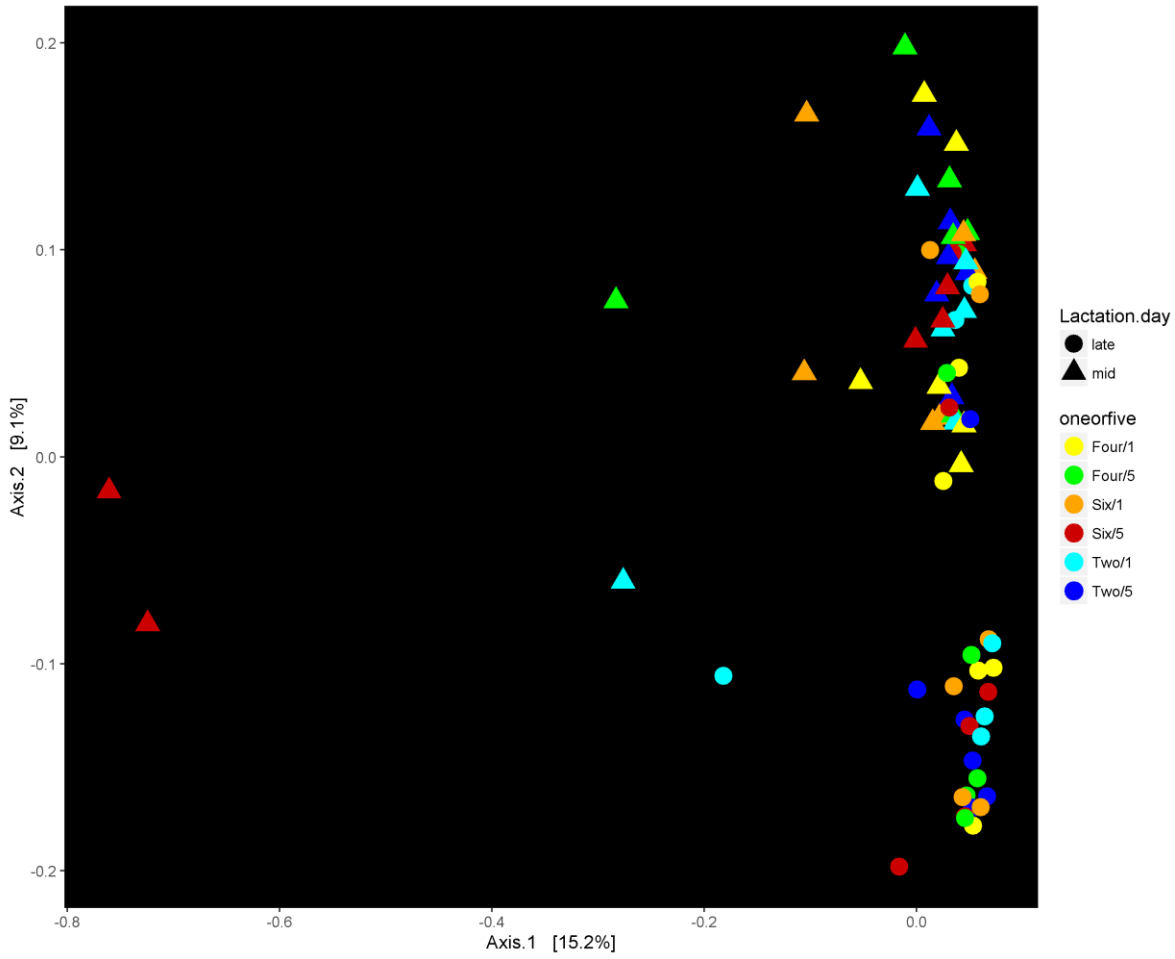


Fig S1: Bray curtis PCoA plots of samples separated by temperature, lactation stage and faceted by lactation stage and day. Four/1=4°C day 1, Four/5= 4°C day 5, Six/1=6°C day 1, Six/5= 6°C day 5, and Two/1=2°C day 1, Two/5= 2°C day 5.

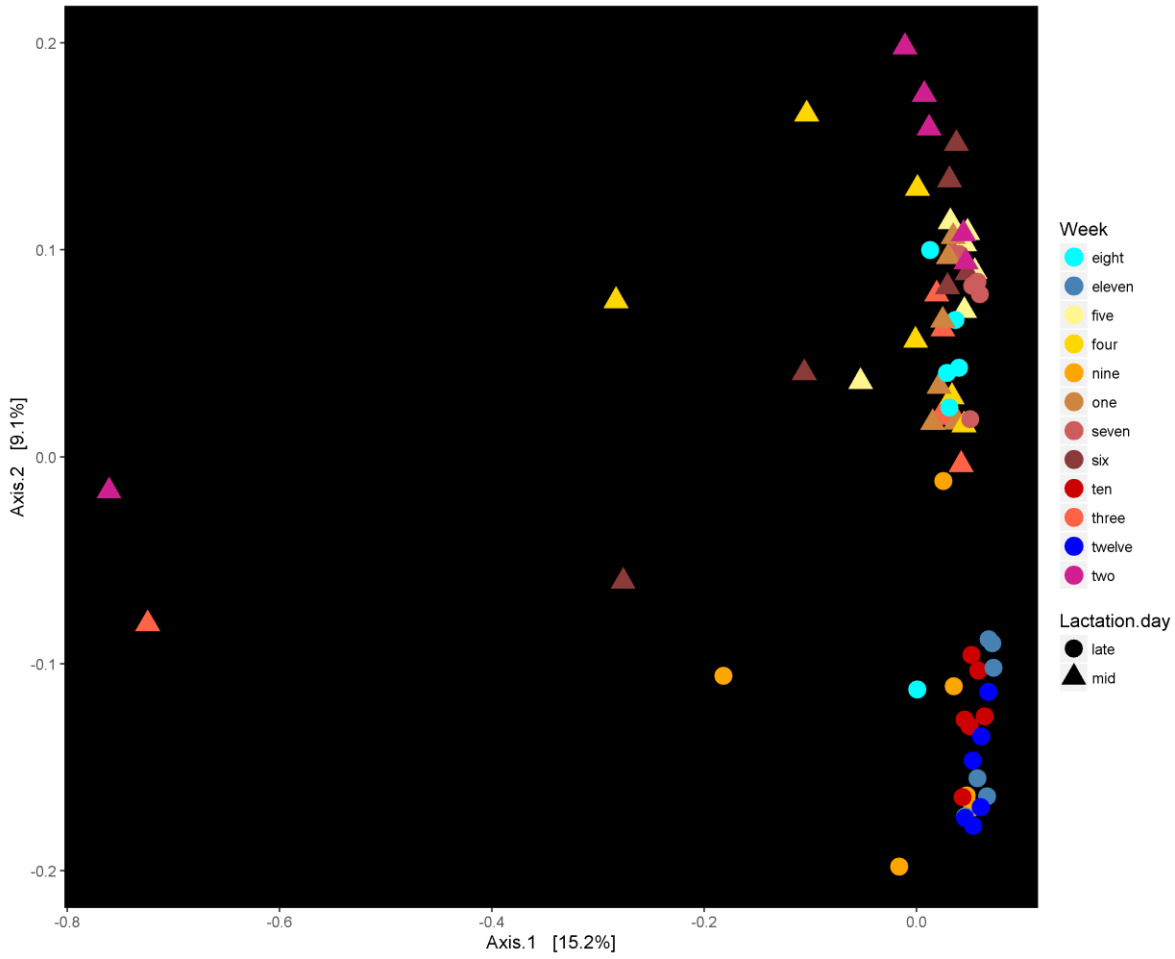


Fig S2: Bray Curtis PCoA plots of samples coloured by week, weeks one to six mid-lactation, weeks seven to twelve late-lactation.

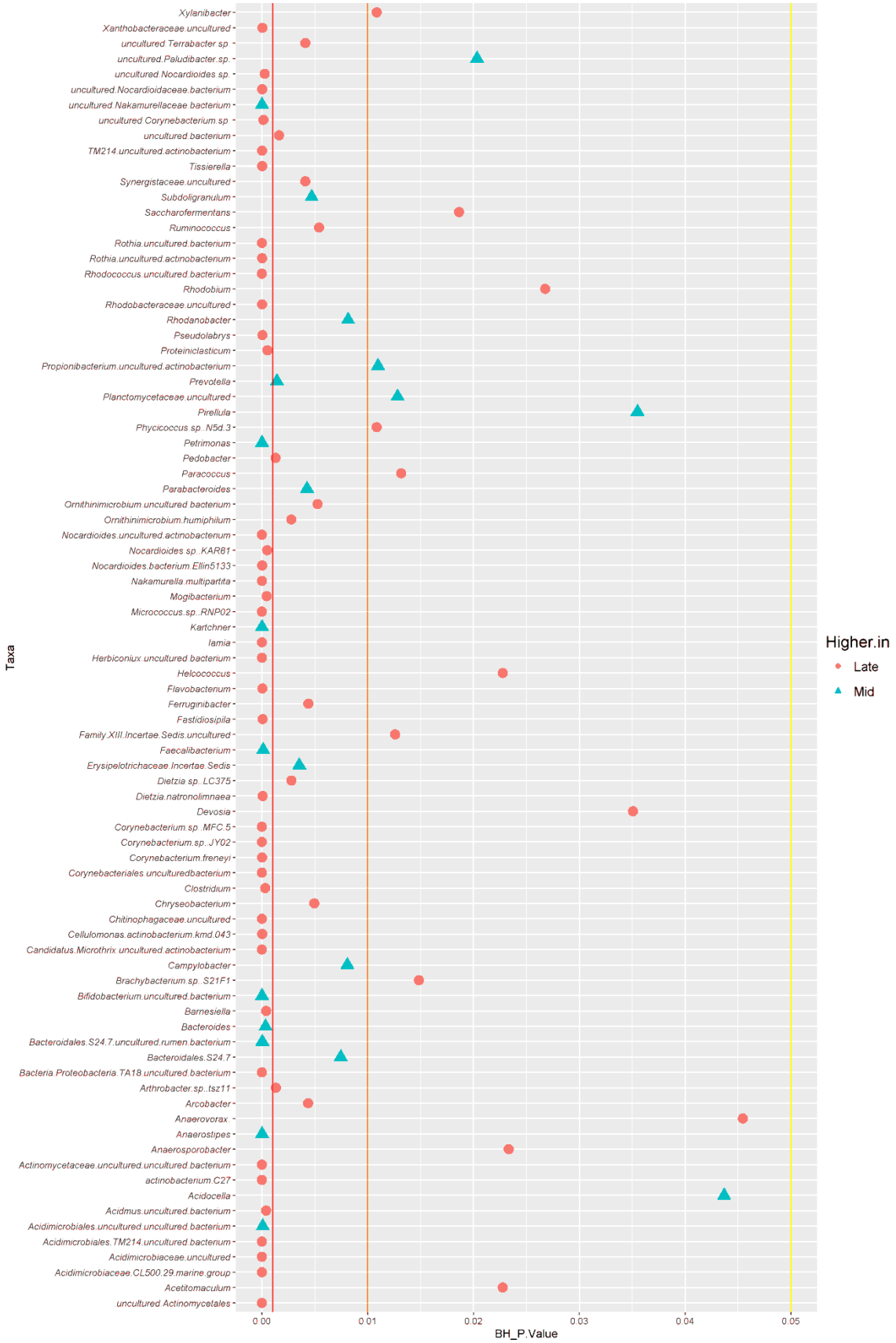
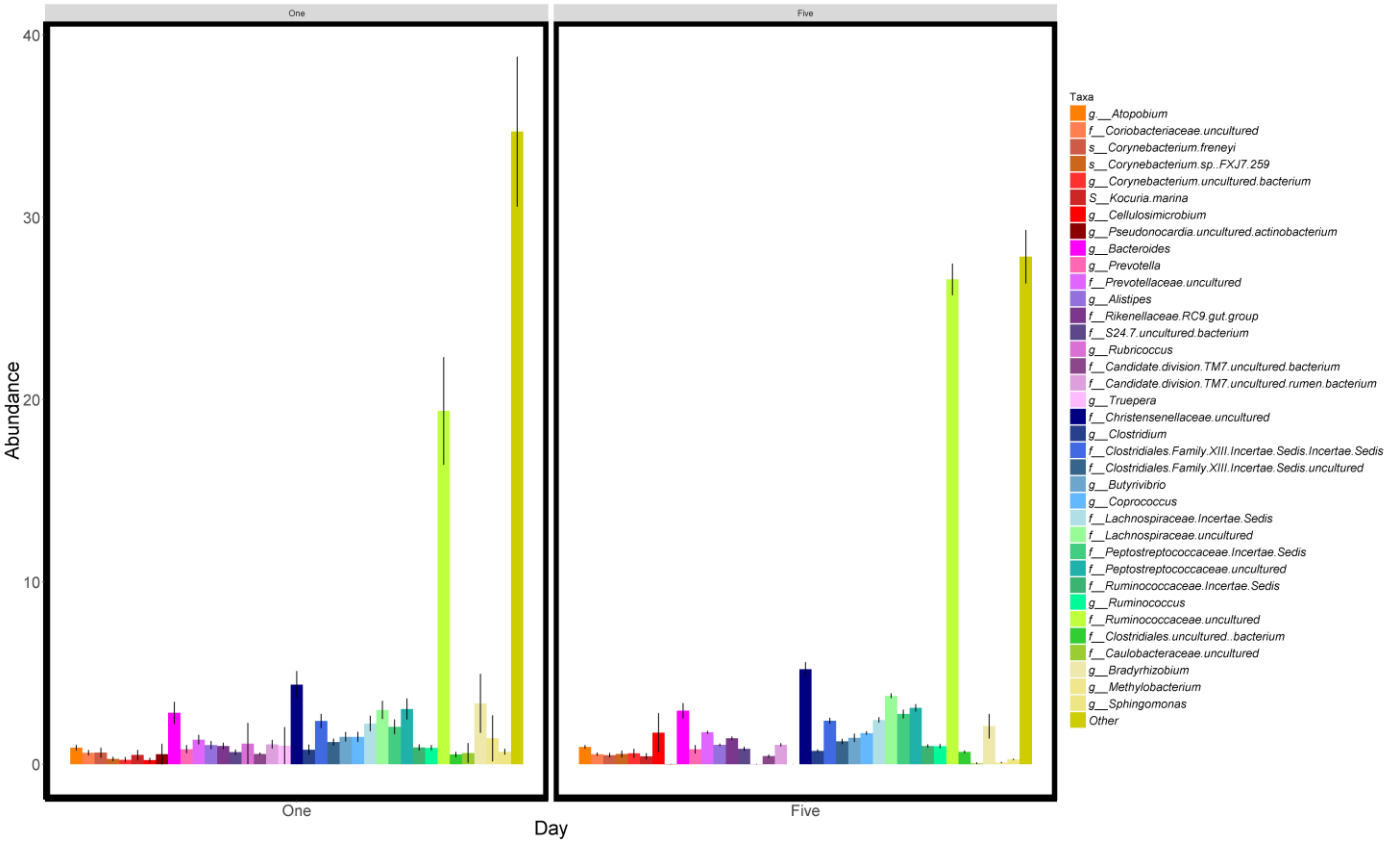


Fig S3: Genera/Species that with significantly differences in abundance between mid and late-lactation after BH correction.



FigS4: Taxa present in 2 ºC samples at day 1 and day 5 in mid-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

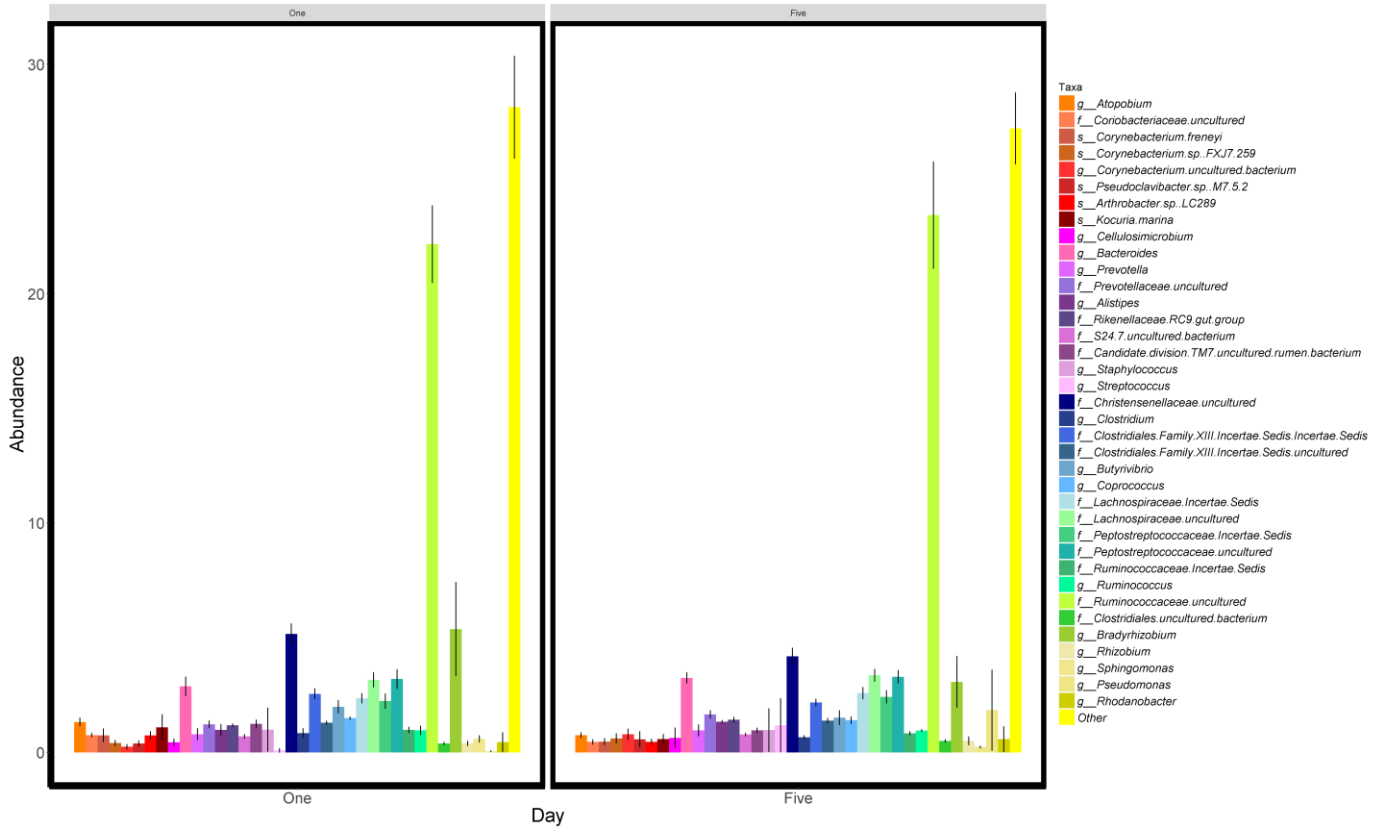


Fig.S5: Taxa present in 4 °C samples at day 1 and day 5 in mid-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

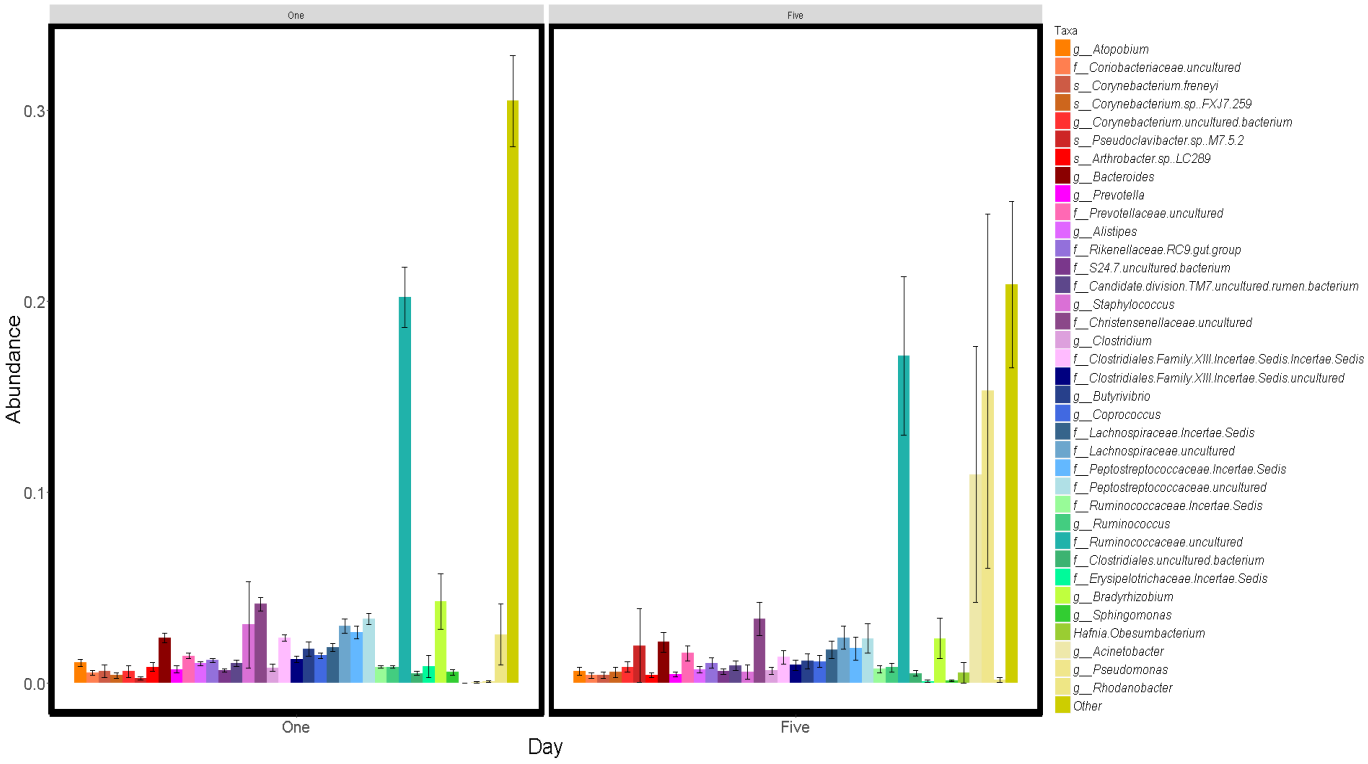


Fig.S6: Taxa present in 6 °C samples at day 1 and day 5 in mid-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

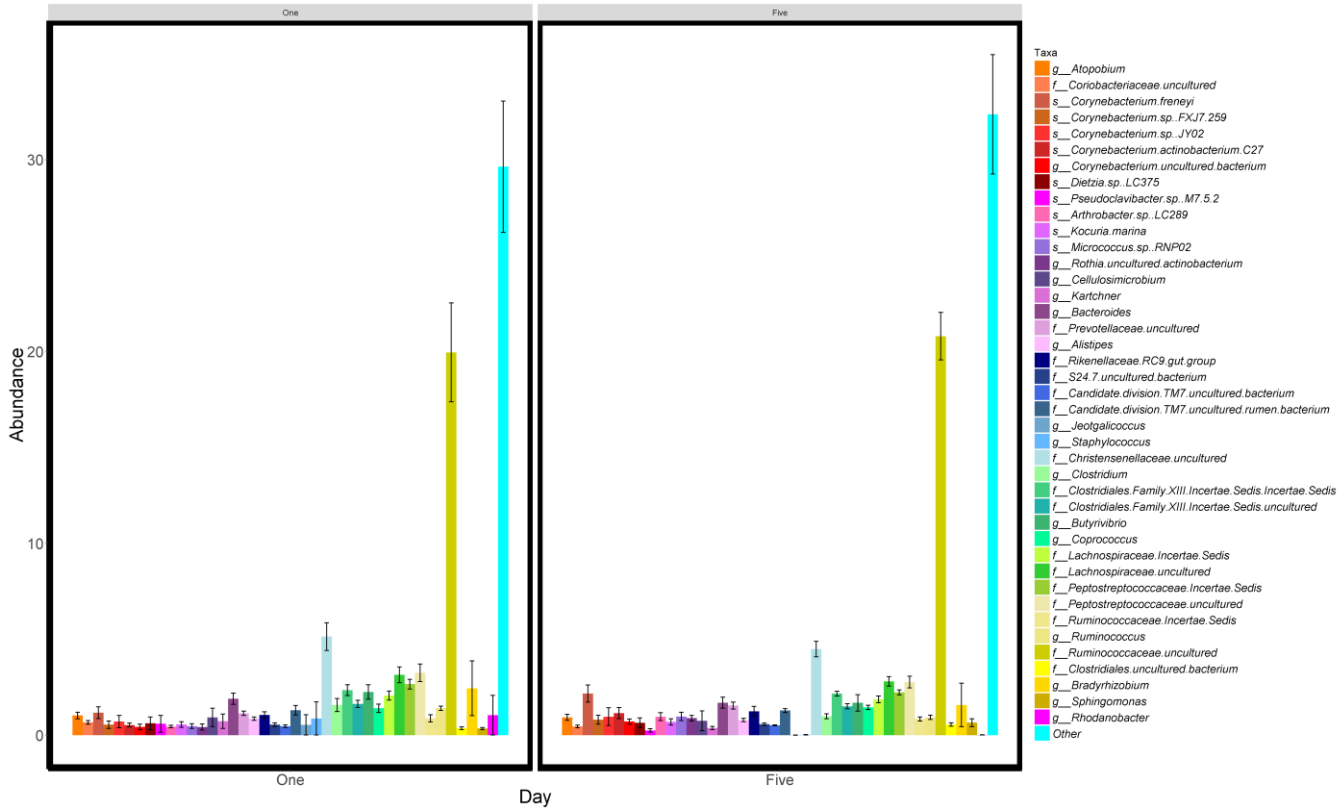


Fig.S7: Taxa present in 2º C samples at day 1 and day 5 in late-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

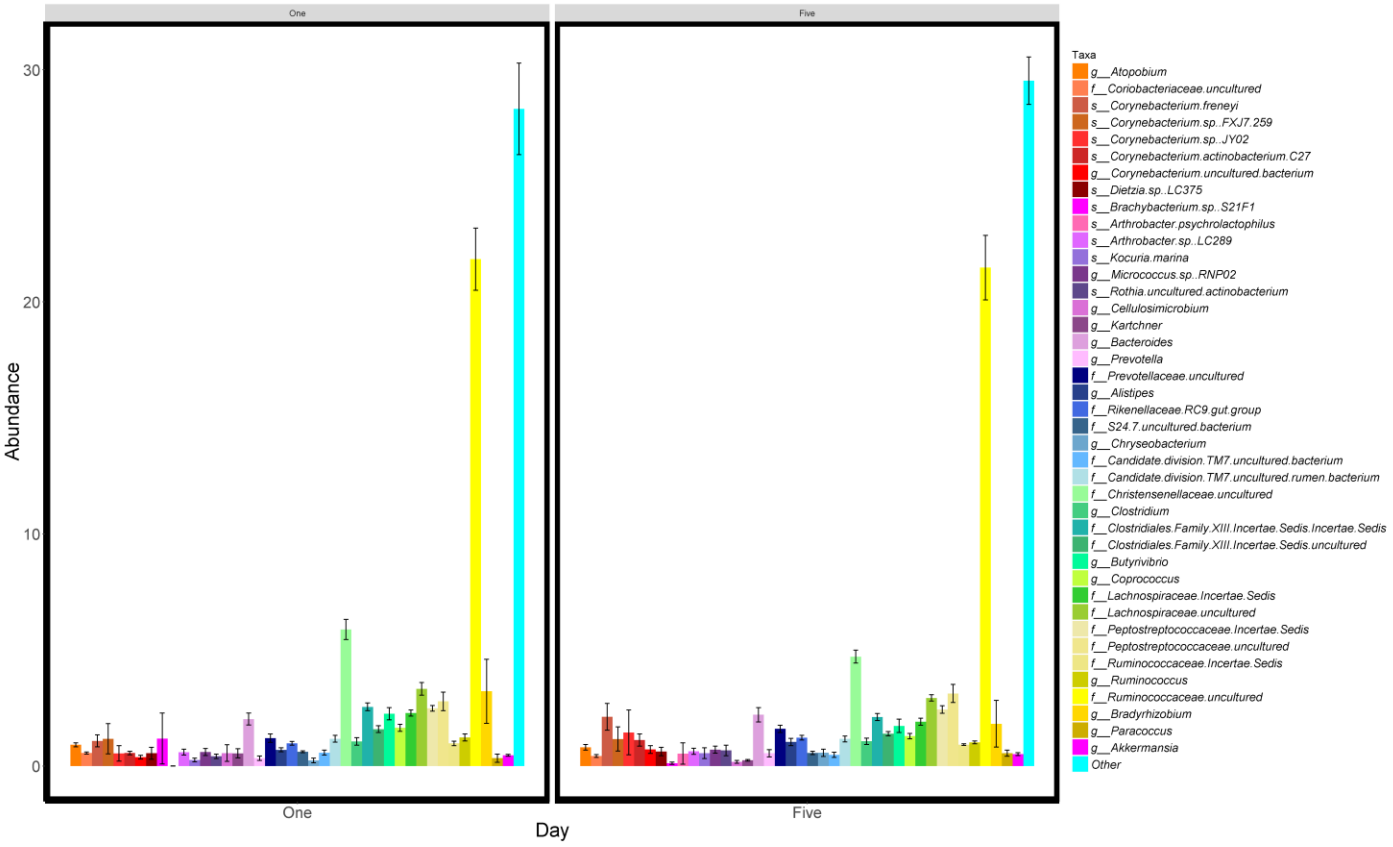


Fig.S8: Taxa present in 4°C samples at day 1 and day 5 in late-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

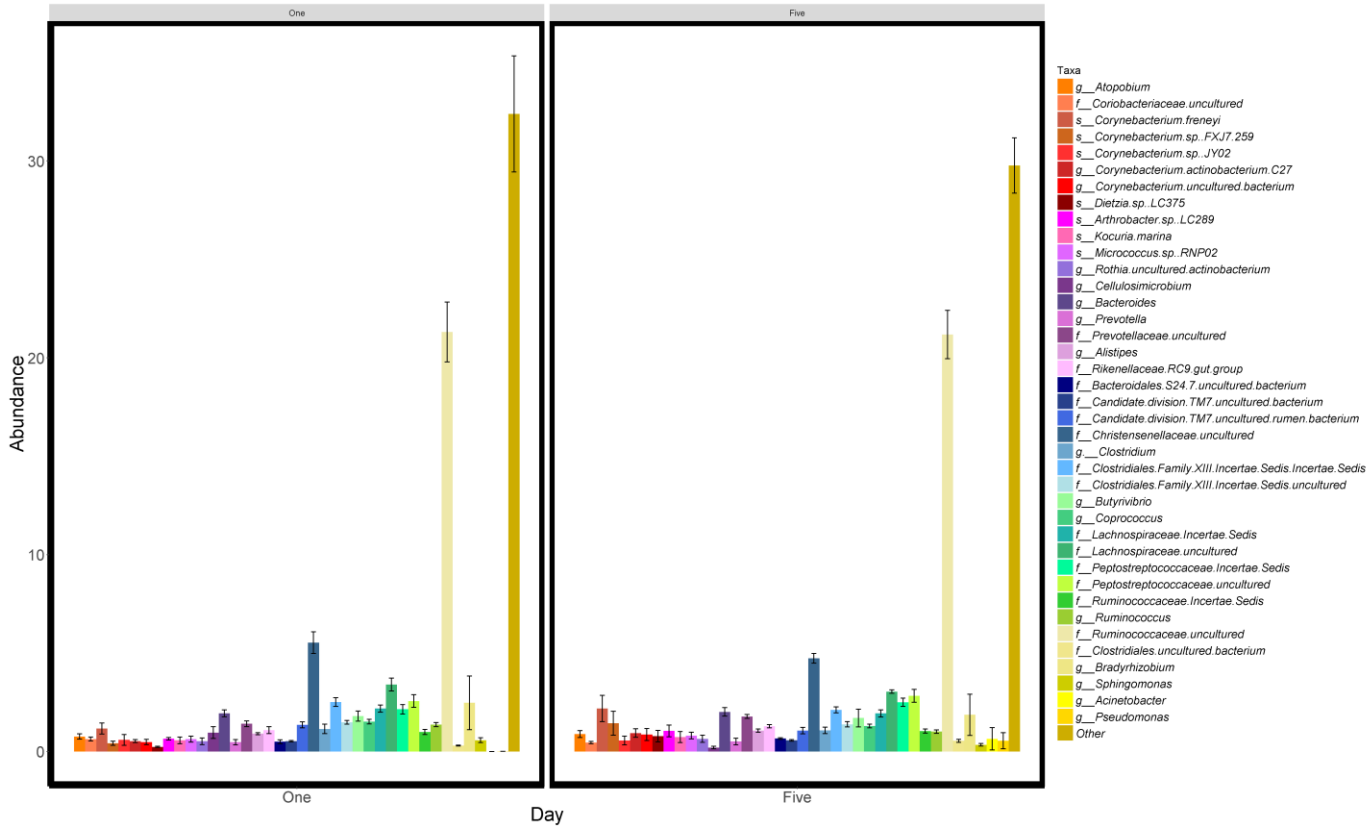


Fig.S9: Taxa present in 6 °C samples at day 1 and day 5 in late-lactation. There were no significant differences in abundance. (f, family; g, genus; s, species).

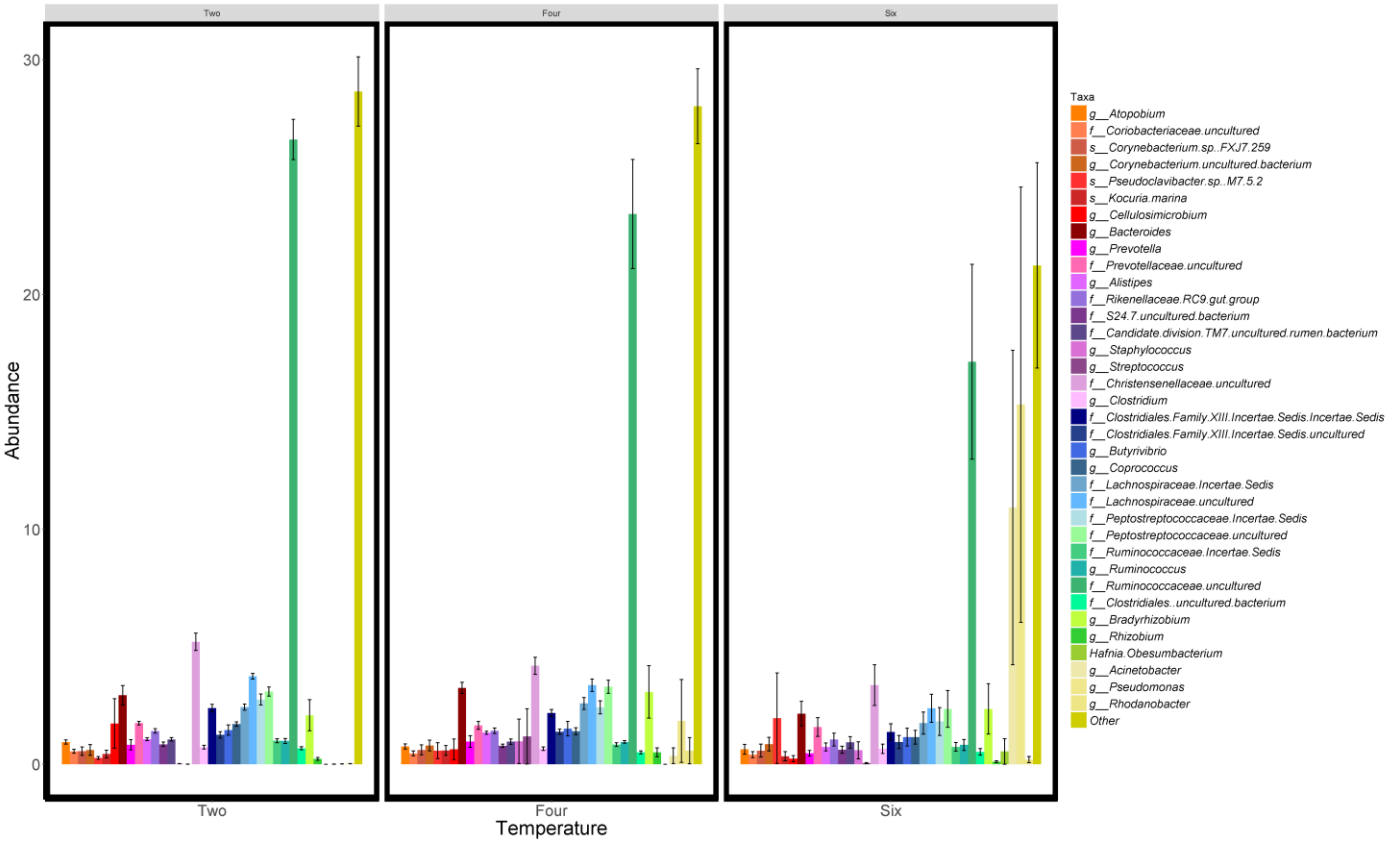


Fig.S10: Taxa present in 2, 4 and 6°C samples in mid-lactation at day 5. There were no significant differences in abundance. (f, family; g, genus; s, species).

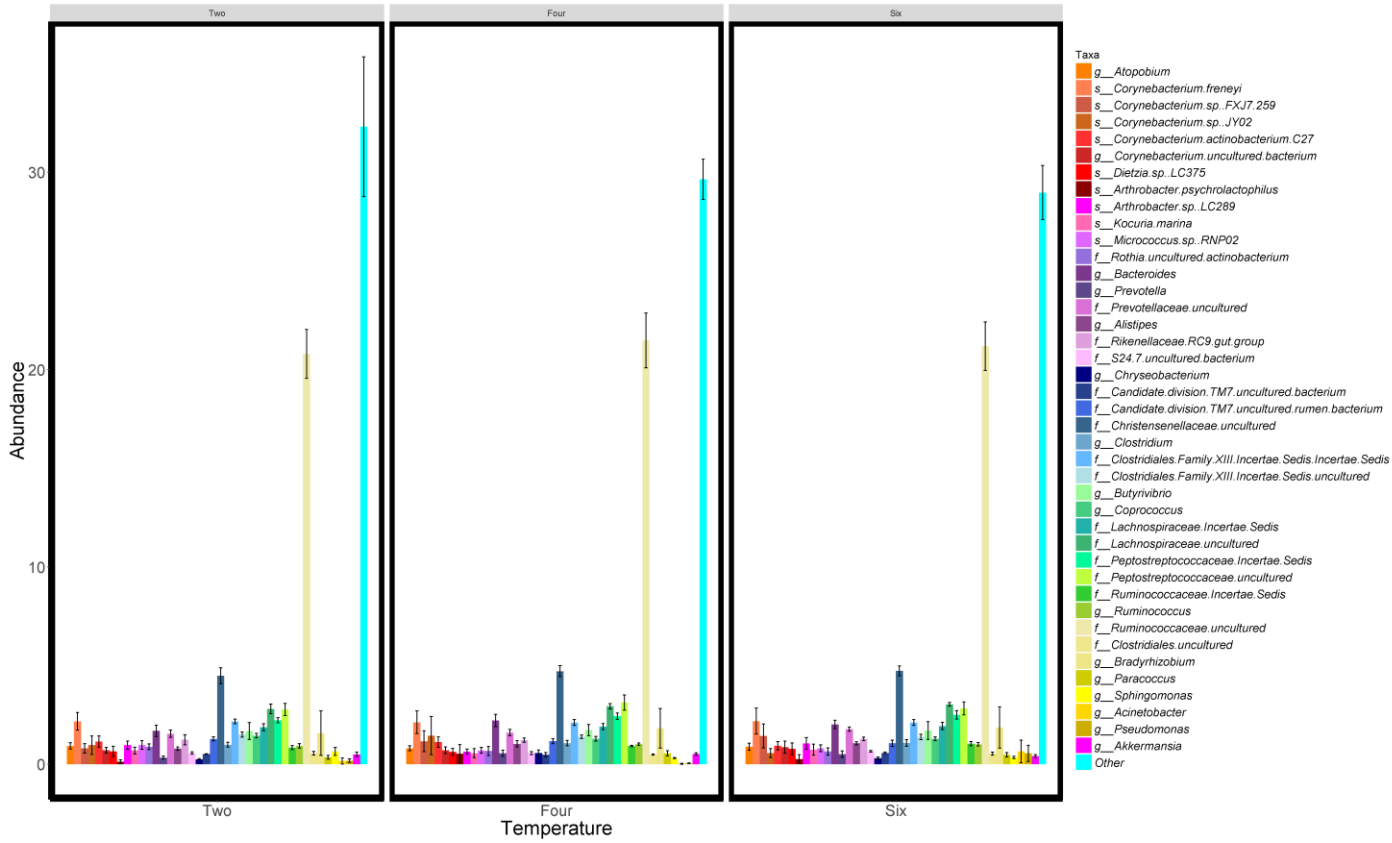


Fig.S11: Taxa present in 2, 4 and 6°C samples late-lactation samples at day 5. There were no significant differences in abundance. (f, family; g, genus; s, species).

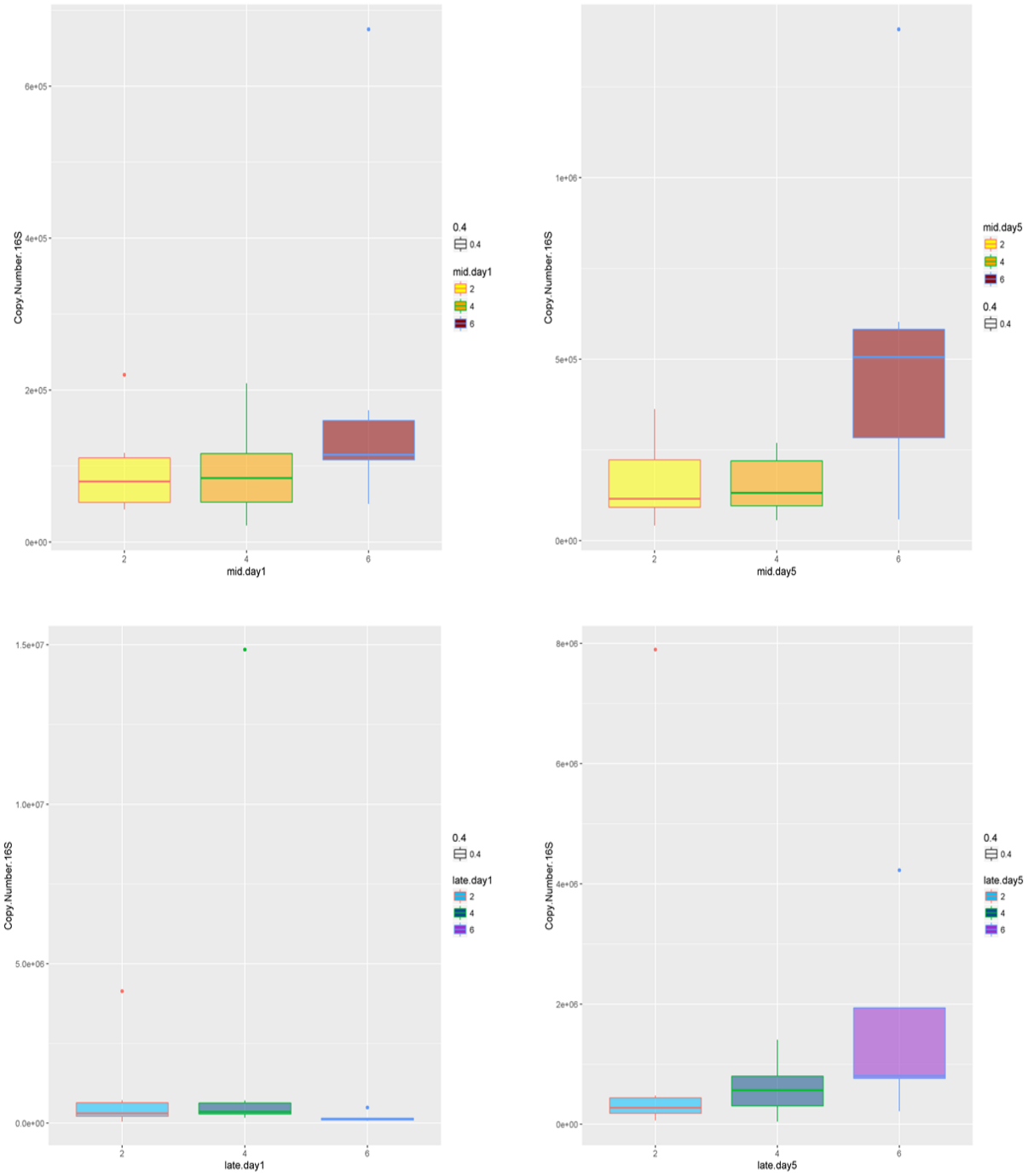


Fig.S12: qPCR total copy number counts for 16S rRNA gene for all temperatures at day 1 and 5 in mid and late-lactation respectively.

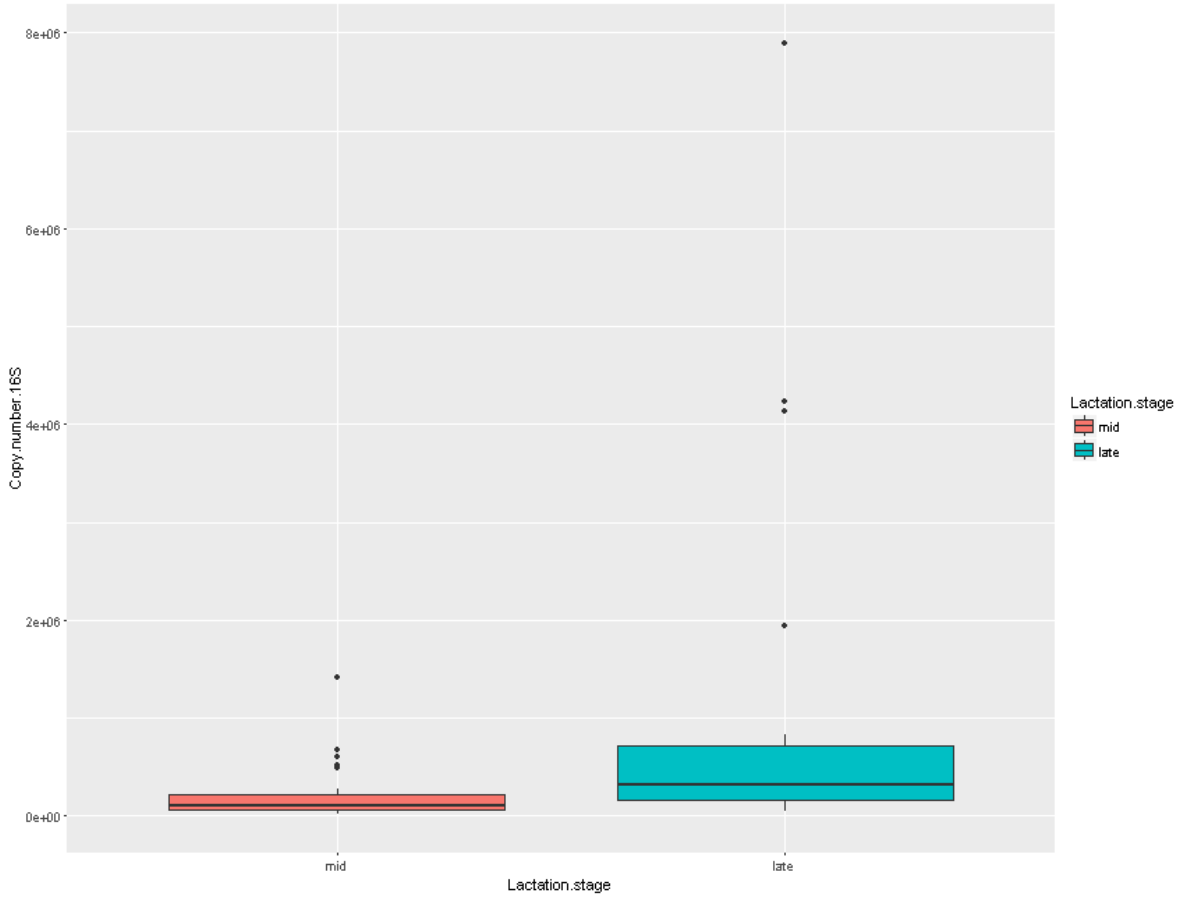


Fig.S13: qPCR total copy number counts for 16S rRNA gene between mid and late lactation.

Chapter 4

Impacts of Seasonal Housing and Teat Preparation on Raw Milk Microbiota: a High-Throughput Sequencing Study

Included as published in Applied and Environmental Microbiology (doi: 10.1128/AEM.02694-16)

4.0 Abstract

In pasture-based systems, changes in dairy herd habitat due to seasonality results in the exposure of animals to different environmental niches. These niches contain distinct microbial communities that may be transferred to raw milk, with potentially important food quality and safety implications for milk producers. It is postulated that the extent to which these microorganisms are transferred could be limited by the inclusion of a teat preparation step prior to milking. Here compositional metagenomics, of a variety of microbial niches on the farm, is employed to study the patterns of microbial movement through the dairy production chain and, in the process, investigate the impact of seasonal housing and inclusion/exclusion of teat preparation regime on the raw milk microbiota from the same herd over two sampling periods, i.e., indoor and outdoor. Beta diversity and network analyses showed that environmental and milk microbiotas separated depending on whether they were sourced from an indoor or outdoor environment. Within these respective habitats, similarities between the milk microbiota and that of teat swab samples and, to a lesser extent, faecal samples were apparent. Indeed, SourceTracker identified the teat surface as the most significant source of contamination, with herd faeces being the next most prevalent source of contamination. In milk from cows grazing outdoors, teat prep significantly increased the numbers of total bacteria present. In summary, sequence-based microbiota analysis identified possible sources of raw milk contamination, and highlighted the influence of environment and farm management practices on the raw milk microbiota.

4.1 Introduction

The impact of the dairy farm environment on the microbial composition of raw milk and raw milk products has been appreciated for some time (Sevi, Albenzio et al. 2003). There are numerous niches that collectively constitute the dairy farm environment and these harbour a vast array of microbes. The transfer of microbes from the farm environment to raw milk can be influenced be a number of factors including farmer hygiene, husbandry practices, herd health, and herd housing (Vacheyrou, Normand et al. 2011). In turn, the microbial composition of raw milk is critically important to its quality, processability and safety.

The microbiota composition of dairy farm niches and of raw milk has typically been examined using traditional plate cultivation-based techniques. These culture-based assays are still widely used by industry and target specific phenotypes, e.g. ability to grow at or survive exposure to particular temperatures (psychrotrophs (Vithanage, Dissanayake et al. 2016), mesophiles (Mhone, Matope et al. 2011), thermodurics (Doyle, Gleeson et al. 2015), or capacity to produce proteases, lipases or other enzymes (Hantsis-Zacharov and Halpern 2007)) or species known to be human pathogens (Vacheyrou, Normand et al. 2011). Using these culture-based techniques, Vacheyrou previously examined possible routes of microbial transfer in farms supplying raw milk for Comte style cheese, revealing that the extent to which milk was contaminated varied depending on the type of barns used to house animals (Vacheyrou, Normand et al. 2011). However, recent advances in molecular microbiology, and in high-throughput DNA sequencing (HTS) in particular, have allowed for a more in-depth analysis of the flow of microbes through environments (Flores, Bates et al. 2011, Knights, Kuczynski et al. 2011, Bokulich and Mills 2013, Bokulich, Ohta et al. 2013, Kembel, Meadow et al. 2014, Bokulich, Bergsveinson et al. 2015).

Indeed, a study of two artisan cheese-making plants observed that spatial diversification within both plants was indicative of “functional adaptations” by microbial communities colonising different fomites within each plant. Spatial diversification between plants confirms the phenomenon of a unique production plant (“house”)-associated microbiota, which was postulated to influence the distinct organoleptic properties of products from each facility (Bokulich and Mills 2013). The facility-specific microbiota developed as a result of the selection pressure introduced by the individual cheese-making processing methods (Bokulich and Mills 2013). The observation of a niche-specific functional adaptation has also been observed in the microbiota of a winery, with the additional observation that the community was influenced by seasonality (Bokulich, Ohta et al. 2013).

The present proof of concept study focuses on the Irish dairy farm system, which is primarily a pasture based system, in which herds are grazed on pasture for the majority of their lactation curve. However, during the winter months, herds are housed indoors. The transition between environments is an important consideration for dairy producers as it is accompanied by changes in exposure to microbes from different niches in the environment as well as dietary changes. Previous, culture-based, efforts to address this question have noted elevated spore counts in bulk tank milk collected from a number of mid-West American farms during summer months on American farms (Buehner, Anand et al. 2014), although elevated numbers of sporeformers can also be an issue when cows are housed indoors if poor quality silage is used (Gleeson, O’Connell et al. 2013). Our study also investigates the impact that teat preparation has on the microbiology of raw milk. This farm management practice has been shown to reduce bacterial counts in milk previously

(Verdier-Metz, Michel et al. 2009) but its impact on the raw milk microbiota has not been reported.

Based on the results of the studies highlighted above, and in the context of the seasonal milk production system applied in Ireland (all cows calved within a 12 week period), it is reasonable to assume that cattle are exposed to niche-specific microbes when housed indoors during winter months, and that these environmental microbes differ significantly from that present when the herd is grazing on pasture during the summer. Such differences would be expected, in turn, to impact on the raw milk microbiota. Specifically, we examined the influence that seasonal housing and grazing conditions have on the microbiota of raw cows' milk. We also examined the influence that the farm management practice of teat preparation (prep) has on the raw milk microbiota in both environments. To address these questions, we applied HTS and a Bayesian inference algorithm to examine environmental sources of bacteria, as well as seasonal changes to the raw milk microbiota driven by changes in habitat.

4.2 Materials and methods

4.2.1 Treatment and Sample collection

Samples were collected from the same herd of Holstein-Friesian dairy cows (n=60) from the Moorepark Research Farm (Fermoy, Co Cork, Ireland) during February (Average days in milk; ADIM= 140) and May (ADIM=200)) of 2015. The milking parlour and equipment were cleaned after each milking as outlined previously (O'Connell, Ruegg et al. 2016). Sampling

phases corresponded to when cows were housed indoors (February) and outdoors on pasture (May). During the indoor sampling period (February) cows were fed grass silage within a cubicle house with automatic scraper cleaning of the central passageway. Cubicle beds were fitted with rubber mats with a daily allowance of ground limestone added to the backend of the cubicle. Cows managed in the outdoor sampling period (May) grazed on perennial ryegrass pasture on a 24h rotational grazing regime. The herd was milked in a 30-unit, 80-degree side-by-side milking parlour (Dairymaster, Causeway, Co Kerry, Ireland). Although most studies incorporating molecular methods focus only on the bulk tank milk (BTM), in this instance, milk from three individual cows was also tested. Three cows with a somatic cell count lower than 100,000 cells/mL were chosen for specific individual sampling before commencement of the study and were used throughout the study. Milk and teat swab samples were collected twice weekly from these three cows throughout the study during the morning milking.

Two pre-milking teat preparation treatments were applied within each sampling phase. One treatment comprised of washing teats with running water, drawing of foremilk, and an application of a pre-milking teat disinfectant (Deosan Teat-foam) (Deosan, Johnson Diversey (Ireland) Ltd, Jamestown RD, Finglas 11, Dublin) followed at least 30 seconds later by drying using individual paper towels, prior to attaching the milking cluster (prep). The second treatment involved no teat preparation prior to cluster attachment for milking (non prep). For both indoor and outdoor sampling periods, the teat treatments applied were as follows: week one, all animals had teats prepped prior to milking; week two, animals were not prepped; week three, teats were prepped prior to milking and week four no teat preparation was carried out. All cows in the herd were subjected to each teat preparation

treatment at each day of sampling. Environmental samples (faeces, bedding, silage grass and surface soil) were collected twice a week on day 1 and day 3, apart from the teat swab samples, which were collected after the teat preparation treatment was applied and prior to cluster attachment for milking on days 2 and 4. Microbial DNA was extracted from all samples using the Powersoil kit (Mobio, Carlsbad CA). Due to the different sample types, the pre-processing protocol for samples varied. At morning's milking on day 2 and 4 of each sampling week, all four teats from the cows were swabbed using one sterile cotton swab per teat (Sarstedt, Ireland). Swabs were dipped in a solution of 3ml of NaCl (0.09%) prior to swabbing to improve recovery (Landers, Hoet et al. 2010). Swabs were drawn across the teat orifice and up the side of each teat avoiding contact with the udder hair. The four swabs from each cow were then pooled in a NaCl solution (12 mL) in a sterile 15 mL falcon tube (Sarstedt, Ireland) and vortexed for 2 minutes. This resulted in one teat pool for each cow sampled at each time point. The pool, including liquid and swab heads, was then centrifuged for 5 minutes at 900 x g to separate the swab heads from the liquid. The supernatant was then removed and transferred to another sterile 15 mL falcon tube. Each pool was then centrifuged at 5444 x g for 30 minutes at 4 ^{o}C. The supernatant was then carefully removed and the resulting pellet was dissolved in the lysis solution from the Powersoil microbead tubes.

Milk samples from the selected three cows were collected within sterilized sampling bottles using the Weighall milk meter on days 2 and 4 of each sampling week (Dairymaster, Causeway, Co Kerry, Ireland). 60 mL of individual milk was used for each extraction. BTM samples representing the complete herd were collected after the morning milking on days 2 and 4. These were collected using 30 mL sterile blue dippa sample tubes (Ocon chemicals).

60 mL of the BTM was used for each extraction. For both individual milk and BTM, milk was aseptically transferred to 15mL Falcon tubes (Sarstedt, Ireland), and centrifuged at 5444 x g for 30 minutes at 4 ^{o}C. The fat layer was carefully removed and the supernatant was decanted. The resulting pellets were then washed using sterile PBS and centrifuged at 14,000 x g for 1 minute. The four pellets for each individual milk and BTM sample were then pooled, to give four samples (three individual milk samples and one BTM sample). Cell pellets were then dissolved in the lysis solution from the microbead tubes from the Powersoil kit.

For faecal pool samples, a pool of the herd's faecal samples was created at each day of sampling. Two faecal pools were collected on each week of sampling on day 1 and 3. To make this pool, equivalent amounts of faecal material were collected from 5 random cow pats and the pool was then homogenised for 2 minutes by vortexing at full speed. DNA was extracted from 250mg of this faecal pool.

Surface soil samples were collected on days 1 and 3 from the paddock from which the herd were grazing. These samples were collected, taking care to avoid collecting faeces or grass using a disposable spatula (VWR, Ireland),250mg of surface soil was used for the soil sample extractions. For bedding, silage and grass samples, 20 g of material was aseptically collected using sterile forceps (VWR, Ireland) and scissors (for grass samples) (Medguard, Co. Meath Ireland) and stored in stomacher bags. For bedding samples 4g of bedding material was collected from 5 cubicles from which the herd had been occupying to create a 20g bedding sample, two bedding samples were collected on each week of the indoor sampling period. For silage samples 20g of silage was collected from where the herd was feeding, two silage samples were collected on each week of the indoor sampling period. For grass samples, 20g

of grass was aseptically collected from the paddock in which the herd had been grazing when outdoors; two grass samples were collected on each week of the outdoor sampling period. Then180 mL of sterile PBS was added to each stomacher bag and the samples were homogenised in a stomacher. The resultant mixture was then aliquoted into 50 mL falcon tubes and centrifuged at 900 x G for 5 minutes to remove solids. Following this, the supernatant was filtered through 0.45 uM nitro cellulose filter membrane (Merck Millipore). After filtration, the membrane was aseptically cut into microbead tubes (Powersoil kit) using a sterile scissors and forceps.

The sample numbers collected included surface soil (n = 8), faeces (n = 16, 8 indoor pools and 8 outdoor pools), silage (n = 8) and bedding (n = 8), as well as teat swabs (n = 48, of which 40 subsequently yielded amplicons - 10 indoor no prep [INP], 11 indoor prep [IP], 11 outdoor prep [OP] and 8 outdoor no prep [ONP]), individual milk samples (n=48, of which 47 subsequently yielded amplicons -12 INP, 12 IP, 11 OP and 12 ONP), bulk tank milk (BTM; n=14, 4 INP, 3 IP, 3 ONP, and 4 OP) and grass (n = 8).

After pre-processing of the samples had been pre-processed and lysis solution added, C1 solution lysis solution (preheated to 60°C) was added to all samples, and followed incubation for 10 minutes at 60°C with vortexing every two minutes for 30 seconds. After this incubation, samples were mechanically lysed at full speed for 10 minutes using a TissueLyser (Qiagen) and then processed as per Powersoil kit protocol. DNA was quantified and quality checked by gel electrophoresis and spectrophotometry on a nanodrop 1000 instrument (Thermo Fisher Scientific Inc).

4.2.2 16S rRNA amplicon sequencing

The V3-V4 variable region of the 16S rRNA gene was amplified from the 149 DNA extracts using the 16S metagenomic sequencing library protocol (Illumina). PCR reactions were completed on the template DNA. Initially, the DNA was amplified with primers specific to the V3-V4 region of the 16S rRNA gene which also incorporates the Illumina overhang adaptor (Forward primer 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG; reverse primer 5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC) (Fouhy, Deane et al. 2015). Each PCR reaction contained DNA template (~10–12ng), 5 µl forward primer (1 µM), 5 µl reverse primer (1 µM), 12.5 µl 2X Kapa HiFi Hotstart ready mix (Anachem, Dublin, Ireland), PCR grade water to a final volume of 25µl. For environmental samples (surface soil, faecal, silage, swabs, bedding, and grass) PCR amplification was carried out as follows: heated lid 110°C, 95°C x 3mins, 25 cycles of 95°C x 30s, 55°C x 30s, 72°C x 30s, then 72°C x 5mins and held at 4°C was used. For milk samples the same cycling parameters were used, accept 32 cycles were used instead of 25 cycles. PCR products were visualised using gel electrophoresis (1X TAE buffer, 1.5% agarose, 100V) and cleaned using AMPure XP magnetic beads (Labplan, Dublin, Ireland). Following this, a subsequent PCR reaction was completed on the purified DNA (5µl) to index each of the samples, allowing samples to be pooled for sequencing on three flow cell and subsequently demultiplexed for analysis. Samples were indexed randomly to prevent any run bias in analysis. Two indexing primers (Illumina Nextera XT indexing primers, Illumina, Sweden) were used per sample. Each PCR reaction contained 5µl index 1 primer (N7xx), 5µl index 2 primer (S5xx), 25µl 2x Kapa HiFi Hot Start Ready mix, 10µl PCR grade water. PCRs were completed as described above, with 8 amplification cycles. PCR products were visualised using gel electrophoresis and

subsequently cleaned (as described above). Samples were quantified using the Qubit (Bio-Sciences, Dublin, Ireland); along with the broad range DNA quantification assay kit (BioSciences) and samples were then pooled in an equimolar fashion. The pooled sample was run on the Agilent Bioanalyser for quality analysis prior to sequencing. The sample pool (4nM) was denatured with 0.2N NaOH, then diluted to 4pM and combined with 10% (v/v) denatured 4pM PhiX, prepared following Illumina guidelines. Samples were sequenced on the MiSeq sequencing platform in the Teagasc sequencing facility, using a 2 x 250 cycle V3 kit, following standard Illumina sequencing protocols.

4.2.3 Bioinformatic and statistical analysis

250 base pair paired-end reads were assembled using FLASH (FLASH: fast length adjustment of short reads to improve genome assemblies) (Magoč and Salzberg 2011). Further processing of paired-end reads including quality filtering based on a quality score of > 25 and removal of mismatched barcodes and sequences below length thresholds was completed using QIIME(Caporaso, Kuczynski et al. 2010). A total of 32,766,563 reads were generated post filtering, with an average of 219,909 per sample. Denoising, chimera detection and clustering into operational taxonomic units (OTUs) (97% identity) were performed using USEARCH v7 (64-bit)(Edgar 2010). OTUs were aligned using PyNAST (python nearest alignment space termination; a flexible tool for aligning sequences to a template alignment) and taxonomy was assigned using BLAST against the SILVA SSURef database release 111. Samples were then rarefied to an even depth of sequences per sample. Alpha diversity was generated in QIIME and the compareGroups function (Subirana, Sanz et al.) was then was then used to determine any statistically significant

differences (P=<0.05) and generate standard deviations between samples based on conditions using the ANOVA test. Beta diversity was calculated in R, using Phyloseq (McMurdie and Holmes 2013) and Bray Curtis distances.Principal coordinate analysis (PCoA) plots were visualised using ggplot2 (Wickham, Chang et al. 2013). Confidence ellipses were generated using stat_ellipse in the ggplot2 package (Wickham, Chang et al. 2013). Network analysis was also carried out using phyloseq and ggplot2. The SourceTracker algorithm (Knights, Kuczynski et al. 2011) was also used to investigate possible sources of environmental contamination in milk from both sampling periods. SourceTracker analysis was carried out at a depth of 13500, with 100 burn-ins and 10 re-starts. The compareGroups function was used in R to compare differences in microbial composition between individual milk, teat swab and faecal pool samples; the Kruskal Wallis test was applied in this instance with Benjamini-Hochberg corrections (Benjamini and Hochberg 1995), to highlight any statistically significant differences (P=<0.05 after correction).

4.2.4 Quantitative PCR

Quantitative PCR (qPCR) was carried out on individual milk samples to determine total bacteria levels in each sample using 16S rRNA gene. qPCR was carried out as described previously (Fouhy, Guinane et al. 2012) except for the use of the equivalent volume of Kappa SYBR fast (Roche Diagnostics) was used instead of SYBR green for the present study. Samples, negative controls (where template DNA was replaced with PCR-grade water) and standards were run in triplicate (technical replicates).

4.3 Results

4.3.1 Microbiota alpha and beta diversity of raw milk, teat surface swabs and environmental samples cluster according to habitat

Samples were collected from the same herd over two sampling periods. Sampling phases corresponded to when the herd was housed indoors and outdoors on pasture, respectively. Across both sampling phases, milk samples were collected from teat prepared (prepped) and non-teat prepped samples. Samples were also classified as either a potential 'source' of microorganisms or a 'sink' (a sample that is liable to contain bacteria originating from a source). Milk samples both from individual cows and BTM were classified as sinks and all environmental samples were classified as sources. After sequencing, the alpha and beta diversity of the bacterial populations present was investigated.

Alpha diversity is the diversity in each sample, using species richness and evenness to calculate the diversity in each environment. There was no significant difference in alpha diversity between the microbiotas of individual indoor and outdoor milk samples from non-prepped animals. Similarly, there was no significant difference in the alpha diversity of the microbiota of indoor milk sourced from animals who underwent teat prep and those that did not. However, the alpha diversity of the outdoor milk microbiota was significantly higher in OP samples relative ONP (P=0.016 Simpsons diversity index, P=0.008 Shannon diversity index; Table 1). A corresponding analysis of the alpha diversity of the microbiota of the teat surface revealed significantly greater diversity (chao1, Shannon, PD whole tree and observed species) among OP samples relative to IP samples (P=<0.01, 0.026, <0.01 and <0.01, respectively; Table 1). No other significant differences in the alpha diversity of teat microbiota samples were observed.

Beta diversity is the diversity between different samples; it provides a measure of dissimilarity between samples. The Bray Curtis Principle Coordinate plot of beta diversity (Fig.1A) depicts all samples from this study with data points coloured by sample origin and shaped according to their designation as source or sink. In this plot it can be observed that samples (soil, grass, bedding, silage, teat surface indoor, teat surface outdoor, faecal indoor pool, faecal outdoor pool, indoor milk, outdoor milk [individual and BTM]) form clusters, which in turn are further separated from one another based on habitat (outdoor/indoor). More specifically, there is a clear separation between samples depending on whether they were collected from an indoor or an outdoor environment. Faeces, teat, individual milk samples and BTM samples also separate based on which environment they were sampled from (indoor/outdoor) (Fig.1A). There are more similarities between samples taken from the same habitat. This includes environmental samples (grass and soil [outdoor] and bedding and silage [indoor]), as seen by the overlaps in the ellipses. Within both habitats, it is apparent that there is an overlap between data points representing the milk sample microbiota and that of teat swab samples, reflecting similarities in their beta diversity (Fig.1A). Teat prep did not result in further sub-clusters within the milk or teat samples (Fig.S1). Faecal pool samples from both habitats separate from one another and are located in relatively close proximity to the corresponding milk and teat samples from the same environment (Fig.1A).

Table.1: Alpha diversity differences between individual milk and teat swab samples.

Milk								
	INP	ONP	IP	OP	P value INP vs ONP	P value IP vs OP	P value INP vs IP	P value ONP vs OP
chao1	3139 (1271)	2733 (833)	3017 (703)	3328 (784)	0.721	0.867	0.99	0.445
Simpson	0.98 (0.02)	0.95 (0.05)	0.98 (0.02)	0.98 (0.02)	0.036	0.885	0.983	0.016
Shannon	8.25 (1.07)	7.49 (1.17)	8.26 (1.07)	9.02 (0.80)	0.309	0.361	1	0.008
PD whole tree	90.3 (29.4)	70.5 (27.2)	93.8 (26.1)	86.3 (23.8)	0.304	0.918	0.99	0.521
observed species	2914 (1232)	2525 (784)	2791 (706)	3036 (752)	0.726	0.922	0.988	0.547

Teat								
	INP	ONP	IP	OP	P value INP vs ONP	P value IP vs OP	P value INP vs IP	P value ONP vs OP
chao1	3373 (792)	4307 (1172)	2949 (536)	4791 (1219)	0.187	<0.001	0.742	0.699
Simpson	0.99 (0.01)	0.99 (0.00)	0.99 (0.01)	0.99 (0.00)	0.99	0.716	0.962	0.997
Shannon	8.54 (0.67)	8.84 (0.41)	8.44 (0.48)	9.17 (0.67)	0.695	0.026	0.977	0.612
PD whole tree	125 (27.0)	157 (37.7)	107 (17.8)	174 (39.9)	0.156	<0.001	0.589	0.665
observed species	3194 (767)	4090 (1119)	2725 (500)	4526 (1188)	0.19	<0.001	0.655	0.741

Numbers in the brackets represent standard deviations. INP= Indoor no prep; ONP= Outdoor no prep; PI= Prep indoor; PO= Prep outdoor

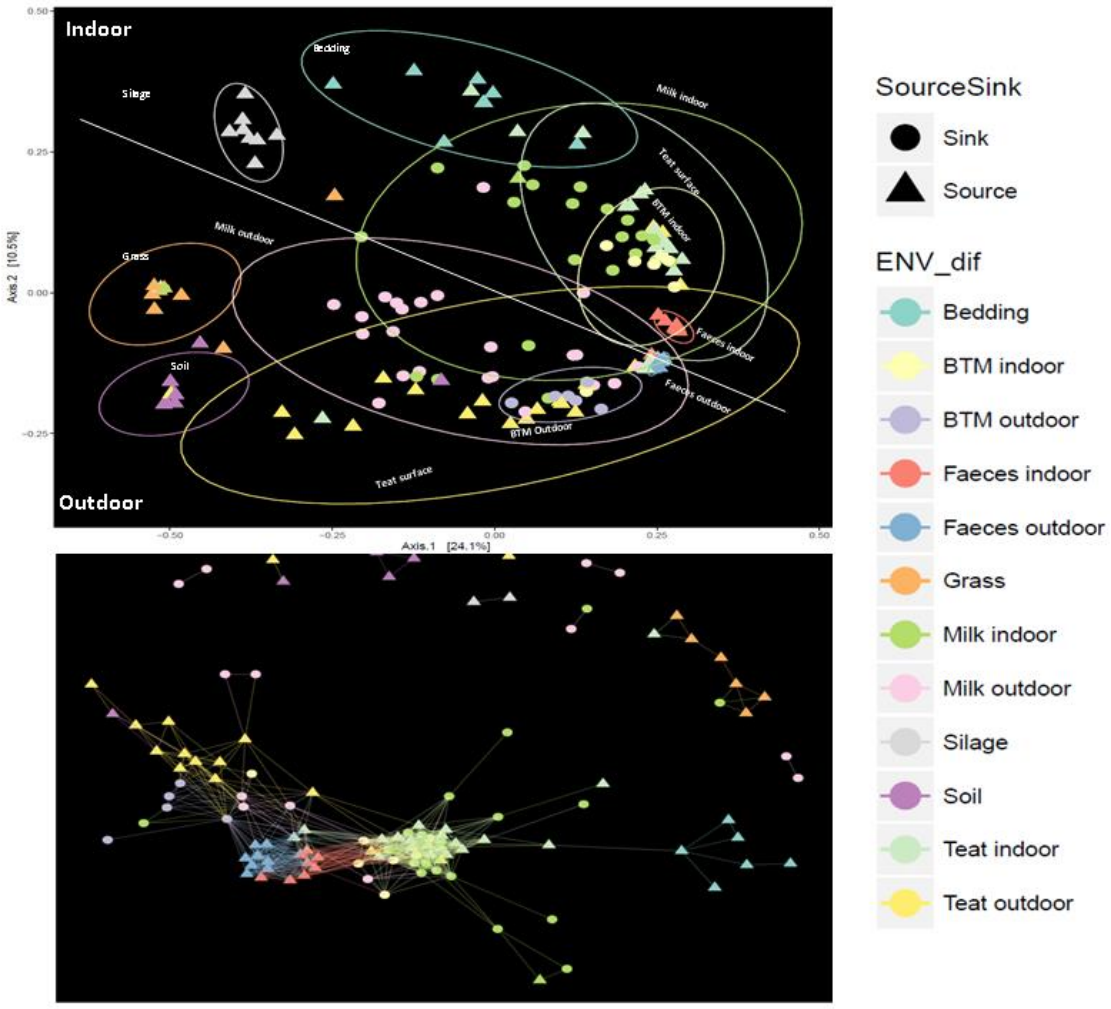


Fig.1: (A)Bray-Curtis PCoA plot of milk and environmental samples, (B) Bray-Curtis Network plot of milk and environmental samples. SourceSink indicates if a sample is classified as a potential source of contamination or a sink for contaminating communities. ENV_dif indicated the sample origin.

4.3.2 Network analysis shows relationships between raw milk and environmental samples

Network plots are a useful graphical tool to illustrate relationships between microbiota datasets. The nodes in this network plot represent samples, and the edges that connect nodes indicate correlations between samples. The network analysis shows relationships that exist between the environmental samples and milk samples (Fig.1B). Consistent with beta diversity data, it is particularly notable that, of the environmental microbiota samples, the faecal pools and teat microbiota are most closely related to the microbiota of the milk samples, thereby identifying faeces and the teat surface as important sources of contamination. These relationships reflect the habitat (indoor or outdoor) from which the samples were collected. There are more edges linking indoor faecal pool samples with indoor BTM samples, than outdoor faecal pool samples with outdoor BTM. Some of the outdoor milk samples are not linked to any of the outdoor sources by edges. This suggests that these niches are not substantial sources of microbial contaminants in these milk samples.

4.3.3 SourceTracker analysis further highlights the contribution of faecal and teat sources to the raw milk microbiota

The SourceTracker model assumes that each individual community (milk, soil, grass, faeces, teat, bedding and silage) is a mixture of communities deposited from other known or unknown source environments and, using a Bayesian approach, the model provides an estimate of the proportion of the community originating from each of the different sources. When a community contains a mixture of taxa that do not match any of the potential source environments studied, that portion of the community is assigned to an "unknown" source. The analysis revealed that the teat surface was the most significant contributor of microbes in milk samples regardless of habitat or teat preparation. Teat surface contaminants constitute a higher proportion of total contaminants in indoor milk compared to outdoor milk, both for individual and for BTM samples. Faeces was the next most important source of contaminants, and had a greater influence on indoor, than outdoor, milk samples, particularly in BTM samples (Fig.2).

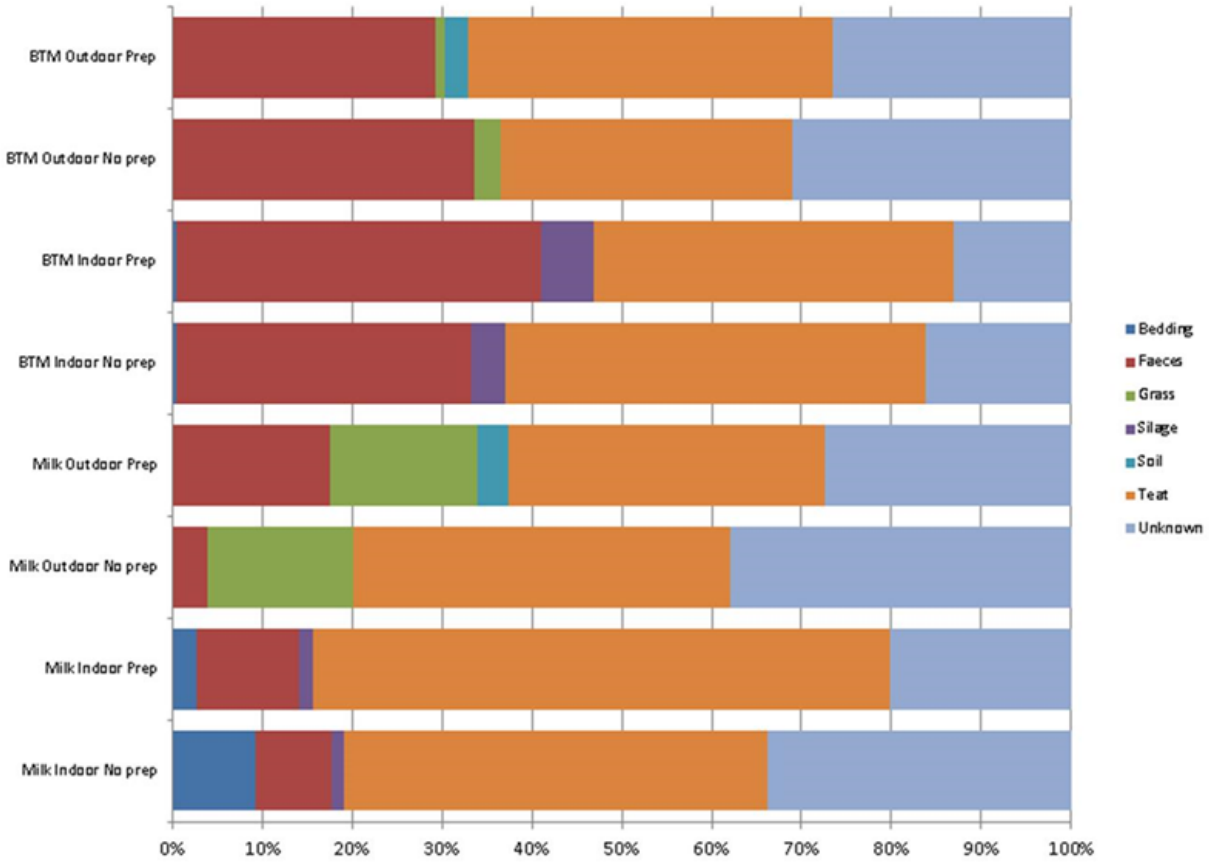


Fig.2: SourceTracker results highlight the percentages of inferred sources of contamination in BTM and individual milk samples.

4.3.4 Taxonomic analysis of raw milk, teat surface and herd faecal microbiota

Graphs representing the microbiota at Family level in the various sample sets are provided in the supplementary data (Fig.S2-3). The compareGroups function was used in R to compare differences in microbial composition between samples. OTUs that differ significantly can be found in the supplementary material (Tables S1-S3). In milk samples from individual animals that did not undergo a teat prep treatment, it was noted that indoor samples contained higher relative proportions of, for example, *Eremococcus*, *Ruminococcus*, *Prevotella*, uncultured *Corynebacteriales* bacterium, and *Ruminococcaceae* Incertae Sedis (P=0.012, 0.012, 0.02, 0.022, 0.028, respectively) and lower proportions of *Pseudomonas*, *Acinetobacter, Lactococcus* and *Tumebacillus* (P=0.003, 0.008, 0.002 and 0.014 respectively), relative to outdoor milk samples. qPCR analysis to determine total bacterial numbers showed that there was significantly more bacteria in indoor milk samples than the equivalent outdoor milk samples (P=0.003) (Table 2). When the corresponding milk samples from individual teat prepped animals were compared, it was noted that 25 genera were present in significantly different proportions in indoor milk samples relative to outdoor-milk samples. Sixteen of these OTUs were higher in indoor samples, these include *Eremococcus*, *Alloiococcus*, *Trichococcus*, *Prevotella*, and *Psychrobacter*, which were all more abundant in indoor samples (P=0.001, 0.001, 0.001, 0.02, and 0.019, respectively). Nine OTUs were higher in PO samples, including *Flavobacterium*, *Sphingomonas* and *Tumebacillus* (P= 0.009, 0.014, and 0.021 respectively). There was no significant difference in total bacterial numbers between the indoor and outdoor milk samples from teat prepped cows (P=0.598) (Fig.3 and Table 2).

The taxonomic data also facilitated an analysis of the specific effects of teat prep on the bacterial composition of the milk produced. In indoor milk samples from individual animals,

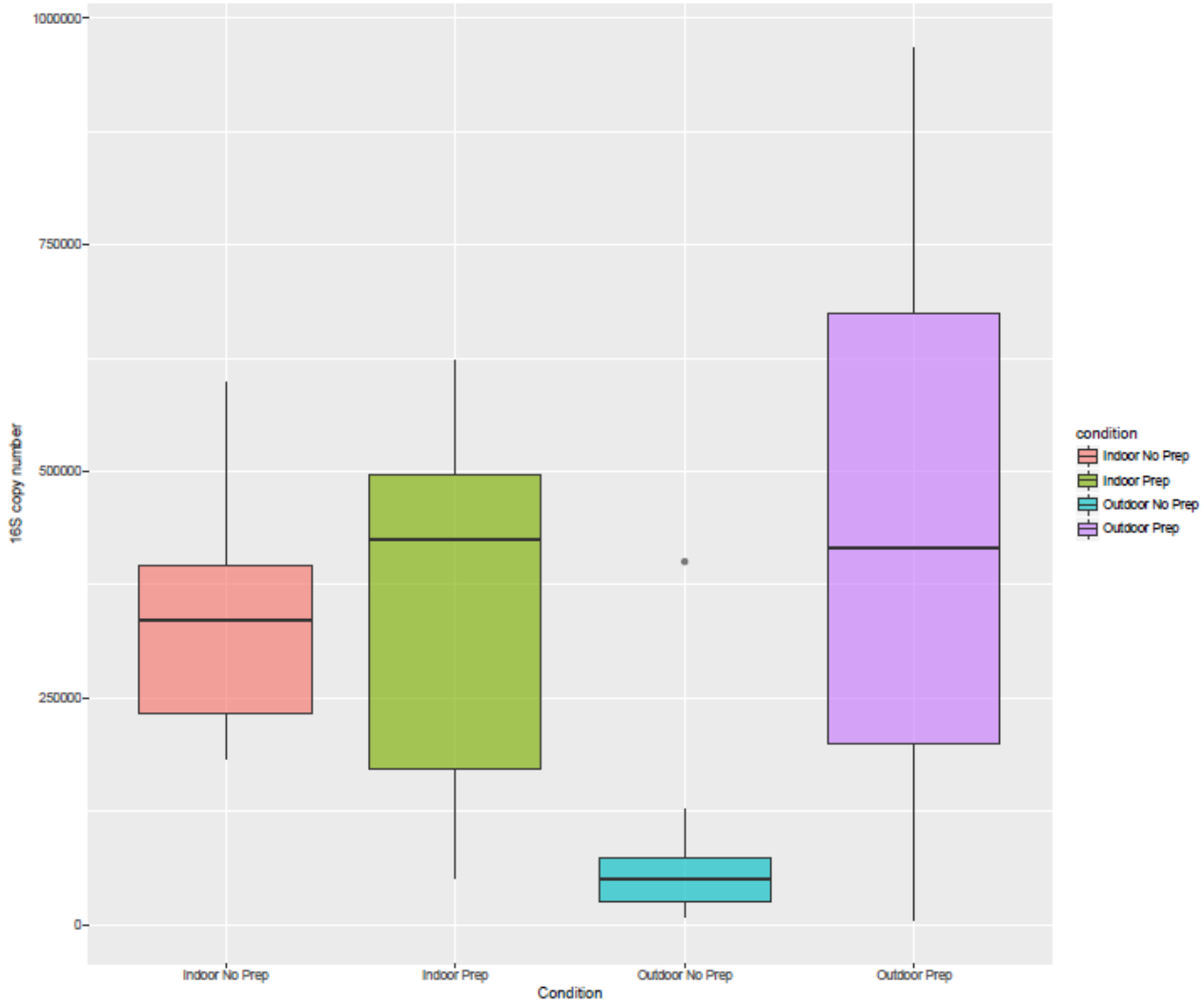


Fig.3: qPCR determination of total bacteria numbers for individual milk samples

Table 2: (A)qPCR determination of total bacteria numbers for individual milk samples, (B) results of comparison total bacterial numbers present in individual milk samples from different conditions.

A	Sample Type	Total bacteria (copies of 16S rRNA gene)
	INP	335500
	IP	424333
	ONP	49600
	OP	416000

B	Comparison	P values
	INP vs IP	0.758
	INP vs ONP	0.003
	IP vs OP	0.598
	ONP vs OP	0.004

INP= Indoor no prep; ONP=Outdoor no prep, PI= Prep indoor, PO= Prep outdoor

it was noted that proportions of *Pseudomonas* were higher in samples from cows which had undergone teat prep (P=0.035) suggesting that, among the indoor teat microbiota, *Pseudomonas* was relatively less sensitive to the antimicrobial effects of the teat prep in indoor samples. qPCR analysis demonstrated that there was no significant difference in total bacterial numbers because of the teat prep (P=0.758) (Table 2). *Pseudomonas*, *Lactococcus* and *Lactobacillus* were among nine genera present in outdoor milk samples that were influenced by teat prep. In the case of the aforementioned genera, proportions were higher in samples when no teat prep was carried out (P=0.011, 0.025, and 0.03, respectively). There were significantly fewer total bacteria in milk samples from non-prepped animals samples compared to samples from prepped animals in the outdoor environment (P=0.004) (Table 2).

The microbiota composition of the teat swabs was also assessed and it was established that, in samples where teat prep did not occur, 18 genera differed significantly in their relative abundance between indoor and outdoor samples. *Trichococcus, Proteiniphilum, and Eremococcus*, as well as *Corynebacterium*, were more abundant in indoor samples (P= 0.012, 0.021, 0.044, and 0.039, respectively) while a further 11 OTU's were present in significantly higher proportions in outdoor samples. In samples where teat preparation was carried out, 60 genera differed significantly between indoor and outdoor samples. Twenty-one of these, including *Eremococcus*, *Proteiniphilum*, *Corynebacterium*, *Psychrobacter Bifidobacterium*, *Trichococcus* and *Prevotella*, were significantly higher in indoor samples (P= 0.001, 0.001 0.002, 0.002 0.003, 0.004, and 0.005, respectively) and thirty-nine genera, including *Stenotrophomonas*, *Xanthomonas* and *Rhizobium*, (P= 0.001, 0.001, and 0.003, respectively) were significantly higher in outdoor samples. Among the outdoor teat samples, there were no significant differences between prepped and non-prepped samples. Among the corresponding indoor teat samples,

proportions of *Variovorax* and *Devosia* were higher in teat samples which were not treated (P=0.033 and 0.043) (Supplementary table 2).

Additionally, it is noteworthy from the stacked bar charts (Fig S1 (B) and (D)) that the composition of individual milk samples differs considerably from that of BTM. More specifically, higher proportions of *Micrococcaceae* and *Flavobacteriaceae* are observed in all individual milk sample types and *Prevotella* and *Rikenellaceae* were higher in BTM samples.

Finally, the availability of faecal pool samples from both the indoor and outdoor environment facilitated a comparison of their composition. At the genus level 15 genera, including *Prevotella*, *Bacteroides* and *Treponema*, were higher in indoor faecal pool samples (P=0.001, 0.002, and 0.021) and a further eight genera, including *Phocaeicola* and *Paludibacter*, were higher in outdoor faecal pool samples (P=0.027 and 0.036) (Supplementary table 3).

4.4 Discussion

The objective of this proof of concept study was to harness the power of next-generation DNA sequencing technologies to investigate the influence that seasonal housing and teat preparation have on the raw milk microbiota from individual cows and in BTM. Furthermore, information potentially revealing the extent to which different microbial niches in the milk production environment influence the microbiota of raw milk was also generated. While, in the past, culture-based investigations to study the source of microorganisms in raw milk have primarily focused on BTM, in this instance samples from a small subset of individual animals was also included. While analysis did not reveal differences between the microbiota alpha diversity of

indoor and outdoor milk samples, beta diversity analysis highlighted a clear separation between samples that are sourced from an indoor versus an outdoor environment. No distinct separation pattern was observed when samples were coloured by teat preparation treatment (Fig S1). Thus, this analysis demonstrates that habitat had a greater impact on the raw milk microbiota than teat preparation.

The SourceTracker algorithm was used as a complementary means of identifying the likely source within the dairy farm environment (soil, silage, bedding, grass, teat, and faeces) of bacteria ultimately found in raw milk and, in the process, also reveals the influence of seasonal housing and farm management practices. Regardless of habitat or treatment, teat surface was again identified as the greatest contributor to the raw milk microbiota, followed by faeces. This is consistent with a previous (culture-based) study, which proposed that the teat skin was a source of microbial populations in raw milk and that farm management and animal grazing practices influenced the diversity and microbiota of raw milk(Verdier-Metz, Gagne et al. 2012).

The taxonomic results also show that habitat had a much greater influence on the raw milk and teat microbiota than teat prep. For instance, in milk samples from cows that were not subjected to teat prep, Gram positive and gut-associated genera were higher in indoor, relative to outdoor milk, such as *Ruminococcus*, *Eremococcus*, *Ruminococcaceae* Incertae Sedis and uncultured *Corynebacteriales* were higher in indoor, relative to outdoor, samples. *Ruminococcus* and *Ruminococcaceae* Incertae Sedis are both gut-associated genera although, from a dairy perspective, *Ruminococcaceae* Incertae Sedis has previously been found in continental type cheese (O'Sullivan, Cotter et al. 2015) and *Ruminococcus* has been detected in raw milk (Quigley, McCarthy et al. 2013) , and in this study these were in higher proportions in INP milk compared to ONP. While, relatively little is known about the uncultured *Corynebacteriales*, the cultured equivalent contains species known to cause mastitis (Hogan, Smith et al. 1988) as well as others

that are found on the surface of surface-ripened cheese (Beresford, Fitzsimons et al. 2001). Similarly, the other genus noted, *Eremococcus*, has not been well characterised, although a typed strain does exist, having been isolated from the vaginal discharge of a thoroughbred horse (Collins, Jovita et al. 1999). Proportions of the Gram negative genus *Prevotella*, which is typically gut-associated was also higher in indoor samples while, for the outdoor samples, the Gram negative genera *Pseudomonas* and *Acinetobacter*, as well as the Gram positive genus *Lactococcus*, were among those that were more dominant. *Pseudomonas* and *Acinetobacter* are both dairy spoilage-associated genera (Hantsis-Zacharov and Halpern 2007) that can have a negative impact on dairy product quality. Lactococci are best known for their positive contribution to the production of fermented dairy products, but can also be isolated from outdoor environments such as grass (Alemayehu, Hannon et al. 2014). These results indicate that indoor milk is more likely to have higher proportions of host/gut associated microbes than outdoor milk while, unsurprisingly, outdoor milk is more likely to contain higher proportions of environmental bacteria.

For milk samples from cows that were teat prepped prior to milking, LAB, such as *Eremococcus*, *Alloiococcus*, and *Trichococcus,* as well as *Psychrobacter*, are also in a significantly higher proportion in IP samples. Interestingly, *Alloiococcus* has not been described in raw milk previously, having instead being associated with human ear infections (Aguirre and Collins 1992). *Trichococcus* has been found in raw milk and dairy waste (Rasolofo, St-Gelais et al. 2010) and *Psychrobacter* have previously been found in teat apexes (Braem, De Vliegher et al. 2012) and in cheese (Quigley, O'Sullivan et al. 2012). Again, in the corresponding OP milk samples soil bacteria such as *Flavobacterium*, *Sphingomonas* and *Tumebacillus* where in higher proportions. This indicates that outdoor milk is more likely to contain increased proportions of soil associated microbes, while indoor milk is more likely to have higher proportions of host/gut bacteria. The

proportions of LAB found in the milk appear to be low in comparison to other studies (Quigley, McCarthy et al. 2013) , this is perhaps due to the protocol used which did not incorporate enzymatic lysis.

In teat swab samples, Gram positive genera such as *Corynebacterium*, *Trichococcus* and *Eremococcus* and Gram negative genera such as *Proteiniphilum* were significantly higher in NPI samples compared to NPO samples. *Proteiniphilum* has previously been associated with the faeces of dairy cattle (Kim and Wells 2016). A number of soil type OTU's were observed to be significantly elevated in NPO, relative to NPI teat swab samples. This indicates that the transmission of soil type bacteria to the teat is greater in periods where cows are grazing outdoors, potentially leading to subsequent transmission from the teat to milk. In teat samples that were prepped, *Corynebacterium*, *Eremococcus* and *Trichococcus* were again more abundant in IP teat samples. *Bifidobacterium* was also present in greater proportions in these samples. Although *Bifidobacterium* is typically associated with the gastrointestinal tract (GIT) of warm blooded mammals (Kim and Wells 2016), it may be significant that prep has previously been shown to cause an increase in Actinobacteria proportions on the teat surface (Verdier-Metz, Michel et al. 2009). With regard to Gram negative bacteria, *Proteiniphilum*, *Psychrobacter* and *Prevotella*, were all significantly more abundant in IP teat swab samples compared to OP samples. In outdoor samples that were teat prepped, many soil type bacteria, including *Rhizobium*, *Xanthomonas*, and *Stenotrophomonas*, were significantly more prevalent compared to OP samples. Thus, soil-type bacteria, also noted on the surface of ONP teat surface, persist even when teat prep occurs.

Using the data generated, it possible to assess the impact of teat preparation on the milk and teat microbiota composition by comparing data from animals that were/were not subjected to a treatment (during the same season). In milk samples, lactic acid bacteria, such as *Lactococcus*

and *Lactobacillus,* and *Pseudomonas* were higher in NPO samples, suggesting that the application teat prep significantly reduced the numbers of these microbes in raw milk. There were no significant differences between PO and NPO teat swab samples. Among indoor teat samples, soil type *Proteobacteria*, such as *Variovorax* and *Devosia*, were more abundant in NPI, relative to PI teats. *Variovorax* has previously been found in hay (Vacheyrou, Normand et al. 2011), and *Devosia* has previously been found in raw milk (Baur, Krewinkel et al. 2015). It was surprising to note that teat prep increased the numbers of total bacteria in both indoor and outdoor milk. Alpha diversity was also found to have increased in milk from cows where teats were prepped prior to milking compared to milk from cows where teat preparation was omitted. It may be that the teat preparation process, including forestripping and drying, weakens the attachment of commensal and contaminating teat canal bacteria and results in their being shed into the milk in greater numbers. This result contrasts findings from culture based analysis on the impact of teat prep on raw milk, which found that it reduced bacterial diversity or counts respectivly (McKinnon, Rowlands et al. 1990, Verdier-Metz, Michel et al. 2009). Further studies will be required to re-examine the influence that teat preparation has on the raw milk microbiota. Another important consideration is that the farm used in this study is a research farm where stringent hygiene practices are upheld. This could perhaps limit the impact that teat preparation has on the raw microbiota

There were considerable differences observed between the individual milk and BTM microbiotas (Fig S1). This may be due to microorganisms in the BTM being acquired from the milking machine and pipes. Indeed, this possibility has been highlighted previously (Quigley, O'Sullivan et al. 2013) but not in the context of DNA-based analysis. Further explorations to definitively establish the basis for these differences is merited.

The availability of faecal microbiota data from multiple samples also facilitated comparative analysis of these samples. It was apparent that the beta diversity of the herd faecal pool microbiota differed significantly from the two sampling periods. From a taxonomic perspective, eight genera were found to be significantly higher in outdoor herd faecal samples and fifteen genera were found to be significantly higher in indoor herd faecal pool samples. *Treponema*, *Prevotella* and *Bacteroides* were among the gut-associated genera that were more prevalent in indoor samples. *Treponema* has previously been associated with digital dermatitis in cattle (Trott, Moeller et al. 2003) and in the bovine rumen (Bekele, Koike et al. 2011). *Phocaeicola* and *Paludibacter* have also been positively associated with valerate in the rumen previously (Mao, Zhang et al. 2012), and were higher in outdoor samples. This difference in faecal microbiota may be influenced by habitat, host physiological changes or by dietary changes associated with the differing habitats. It is also possible that transmission of bacteria from faecal origin may differ based on habitat due to the differences in the microbiota seen here.

Here, high-throughput DNA sequencing has facilitated the analysis of the microbiota of raw milk samples in parallel with samples from the dairy farm environment. The results provide a more detailed insight into the composition of these microbial populations while also allowing an examination of the relationship between the microbiota of these environments and of raw milk. This analysis highlights that herd habitat is a significant driver for milk microbiota composition, and that teat prep has a much more limited impact on the raw milk microbiota. In the process it is made apparent that high-throughput sequencing can be an extremely insightful tool to help better understand the movement of microbes from the environment into the food chain.

5.5 References

Aguirre, M. and M. Collins (1992). "Phylogenetic analysis of Alloiococcus otitis gen. nov., sp. nov., an organism from human middle ear fluid." International Journal of Systematic and Evolutionary Microbiology **42**(1): 79-83.

Alemayehu, D., J. A. Hannon, O. McAuliffe and R. P. Ross (2014). "Characterization of plant-derived lactococci on the basis of their volatile compounds profile when grown in milk." International journal of food microbiology **172**: 57-61.

Baur, C., M. Krewinkel, B. Kranz, M. von Neubeck, M. Wenning, S. Scherer, M. Stoeckel, J. Hinrichs, T. Stressler and L. Fischer (2015). "Quantification of the proteolytic and lipolytic activity of microorganisms isolated from raw milk." International Dairy Journal **49**: 23-29.

Bekele, A. Z., S. Koike and Y. Kobayashi (2011). "Phylogenetic diversity and dietary association of rumen Treponema revealed using group-specific 16S rRNA gene-based analysis." FEMS microbiology letters **316**(1): 51-60.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society. Series B (Methodological): 289-300.

Beresford, T. P., N. A. Fitzsimons, N. L. Brennan and T. M. Cogan (2001). "Recent advances in cheese microbiology." International Dairy Journal **11**(4): 259-274.

Bokulich, N. A., J. Bergsveinson, B. Ziola and D. A. Mills (2015). "Mapping microbial ecosystems and spoilage-gene flow in breweries highlights patterns of contamination and resistance." eLife **4**: e04634.

Bokulich, N. A. and D. A. Mills (2013). "Facility-specific “house” microbiome drives microbial landscapes of artisan cheesemaking plants." Applied and environmental microbiology **79**(17): 5214-5223.

Bokulich, N. A., M. Ohta, P. M. Richardson and D. A. Mills (2013). "Monitoring seasonal changes in winery-resident microbiota." PloS one **8**(6): e66437.

Braem, G., S. De Vliegher, B. Verbist, M. Heyndrickx, F. Leroy and L. De Vuyst (2012). "Culture-independent exploration of the teat apex microbiota of dairy cows reveals a wide bacterial species diversity." Veterinary microbiology **157**(3): 383-390.

Buehner, K. P., S. Anand and A. Garcia (2014). "Prevalence of thermoduric bacteria and spores on 10 Midwest dairy farms." Journal of dairy science **97**(11): 6777-6784.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich and J. I. Gordon (2010). "QIIME allows analysis of high-throughput community sequencing data." Nature methods **7**(5): 335-336.

Collins, M. D., M. R. Jovita, P. A. Lawson, E. Falsen and G. Foster (1999). "Characterization of a novel Gram-positive, catalase-negative coccus from horses: description of Eremococcus coleocola gen. nov. sp. nov." International Journal of Systematic and Evolutionary Microbiology **49**(4): 1381-1385.

Doyle, C. J., D. Gleeson, K. Jordan, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." International journal of food microbiology **197**: 77-87.

Edgar, R. C. (2010). "Search and clustering orders of magnitude faster than BLAST." Bioinformatics **26**(19): 2460-2461.

Flores, G. E., S. T. Bates, D. Knights, C. L. Lauber, J. Stombaugh, R. Knight and N. Fierer (2011). "Microbial biogeography of public restroom surfaces." PLoS One **6**(11): e28132.

Fouhy, F., J. Deane, M. C. Rea, Ó. O'Sullivan, R. P. Ross, G. O'Callaghan, B. J. Plant and C. Stanton (2015). "The Effects of Freezing on Faecal Microbiota as Determined Using MiSeq Sequencing and Culture-Based Investigations." PloS one **10**(3): e0119355.

Fouhy, F., C. M. Guinane, S. Hussey, R. Wall, C. A. Ryan, E. M. Dempsey, B. Murphy, R. P. Ross, G. F. Fitzgerald, C. Stanton and P. D. Cotter (2012). "High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin." Antimicrobial agents and chemotherapy **56**(11): 5811-5820.

Gleeson, D., A. O'Connell and K. Jordan (2013). "Review of potential sources and control of thermoduric bacteria in bulk-tank milk." Irish Journal of Agricultural and Food Research **52**(2): 217-227.

Hantsis-Zacharov, E. and M. Halpern (2007). "Culturable psychrotrophic bacterial communities in raw milk and their proteolytic and lipolytic traits." Applied and environmental microbiology **73**(22): 7162-7168.

Hogan, J., K. Smith, D. Todhunter and P. Schoenberger (1988). "Rate of environmental mastitis in quarters infected with Corynebacterium bovis and Staphylococcus species." Journal of dairy science **71**(9): 2520-2525.

Kembel, S. W., J. F. Meadow, T. K. O'Connor, G. Mhuireach, D. Northcutt, J. Kline, M. Moriyama, G. Brown, B. J. Bohannan and J. L. Green (2014). "Architectural design drives the biogeography of indoor bacterial communities." PloS one **9**(1): e87093.

Kim, M. and J. E. Wells (2016). "A Meta-analysis of Bacterial Diversity in the Feces of Cattle." Current microbiology **72**(2): 145-151.

Knights, D., J. Kuczynski, E. S. Charlson, J. Zaneveld, M. C. Mozer, R. G. Collman, F. D. Bushman, R. Knight and S. T. Kelley (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature methods **8**(9): 761-763.

Landers, T. F., A. Hoet and T. E. Wittum (2010). "Swab type, moistening, and preenrichment for Staphylococcus aureus on environmental surfaces." Journal of clinical microbiology **48**(6): 2235-2236.

Magoč, T. and S. L. Salzberg (2011). "FLASH: fast length adjustment of short reads to improve genome assemblies." Bioinformatics **27**(21): 2957-2963.

Mao, S., R. Zhang, D. Wang and W. Zhu (2012). "The diversity of the fecal bacterial community and its relationship with the concentration of volatile fatty acids in the feces during subacute rumen acidosis in dairy cows." BMC Veterinary Research **8**(1): 1-13.

McKinnon, C. H., G. J. Rowlands and A. J. Bramley (1990). "The effect of udder preparation before milking and contamination from the milking plant on bacterial numbers in bulk milk of eight dairy herds." Journal of Dairy Research **57**(03): 307-318.

McMurdie, P. J. and S. Holmes (2013). "phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data."

Mhone, T. A., G. Matope and P. T. Saidi (2011). "Aerobic bacterial, coliform, Escherichia coli and Staphylococcus aureus counts of raw and processed milk from selected smallholder dairy farms of Zimbabwe." International Journal of Food Microbiology **151**(2): 223-228.

O'Sullivan, D. J., P. D. Cotter, O. O'Sullivan, L. Giblin, P. L. McSweeney and J. J. Sheehan (2015). "Temporal and spatial differences in microbial composition during the manufacture of a Continental-type cheese." Applied and environmental microbiology **81**(7): 2525-2533.

O'Connell, A., P. Ruegg, K. Jordan, B. O'Brien and D. Gleeson (2016). "The effect of storage temperature and duration on the microbial quality of bulk tank milk." Journal of dairy science **99**(5): 3367-3374.

Quigley, L., R. McCarthy, O. O'Sullivan, T. P. Beresford, G. F. Fitzgerald, R. P. Ross, C. Stanton and P. D. Cotter (2013). "The microbial content of raw and pasteurized cow's milk as determined by molecular approaches." Journal of Dairy Science.

Quigley, L., O. O'Sullivan, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2012). "High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses." Applied and environmental microbiology **78**(16): 5717-5723.

Quigley, L., O. O'Sullivan, C. Stanton, T. P. Beresford, R. P. Ross, G. F. Fitzgerald and P. D. Cotter (2013). "The complex microbiota of raw milk." FEMS microbiology reviews **37**(5): 664-698.

Rasolofo, E. A., D. St-Gelais, G. LaPointe and D. Roy (2010). "Molecular analysis of bacterial population structure and dynamics during cold storage of untreated and treated milk." International journal of food microbiology **138**(1): 108-118.

Sevi, A., M. Albenzio, A. Muscio, D. Casamassima and P. Centoducati (2003). "Effects of litter management on airborne particulates in sheep houses and on the yield and quality of ewe milk." Livestock production science **81**(1): 1-9.

Subirana, I. V., J. Sanz, H. Lucas, G. Peñafiel and J. Giménez D.(2013). compareGroups: Descriptive analysis by groups: R package version 2.0. Retrieved September 2013.

Trott, D. J., M. R. Moeller, R. L. Zuerner, J. P. Goff, W. R. Waters, D. P. Alt, R. L. Walker and M. J. Wannemuehler (2003). "Characterization of Treponema phagedenis-like spirochetes isolated from papillomatous digital dermatitis lesions in dairy cattle." Journal of clinical microbiology **41**(6): 2522-2529.

Vacheyrou, M., A.-C. Normand, P. Guyot, C. Cassagne, R. Piarroux and Y. Bouton (2011). "Cultivable microbial communities in raw cow milk and potential transfers from stables of sixteen French farms." International journal of food microbiology **146**(3): 253-262.

Verdier-Metz, I., G. Gagne, S. Bornes, F. Monsallier, P. Veisseire, C. Delbès-Paus and M.-C. Montel (2012). "Cow teat skin, a potential source of diverse microbial populations for cheese production." Applied and environmental microbiology **78**(2): 326-333.

Verdier-Metz, I., V. Michel, C. Delbes and M.-C. Montel (2009). "Do milking practices influence the bacterial diversity of raw milk?" Food Microbiology **26**(3): 305-310.

Vithanage, N. R., M. Dissanayake, G. Bolge, E. A. Palombo, T. R. Yeager and N. Datta (2016). "Biodiversity of culturable psychrotrophic microbiota in raw milk attributable to refrigeration conditions, seasonality and their spoilage potential." International Dairy Journal **57**: 80–90.

Wickham, H., W. Chang and M. H. Wickham (2013). "Package 'ggplot2'."

Supplementary Material

Table S1: Corrected P values for Kruskal Wallis test on individual milk samples.

Milk		
NPI vs NPO	**P value**	**Higher in**
Lactococcus	0.002	NPO
Pseudomonas	0.003	NPO
Acinetobacter	0.008	NPO
Eremococcus	0.012	NPI
Ruminococcus	0.012	NPI
Tumebacillus	0.014	NPO
Prevotella	0.02	NPI
Corynebacteriales uncultured bacterium	0.022	NPI
Ruminococcaceae Incertae Sedis	0.028	NPI
Bacteroidales uncultured bacterium	0.049	NPI

NPI vsPI	**P value**	**Higher in**
Pseudomonas	0.035	PI

NPO vs PO	**P value**	**Higher in**
uncultured *Verrucomicrobia* bacterium	0.005	PO
Exiguobacterium	0.009	PO
Pseudomonas	0.011	NPO
DA101 soil group uncultured bacterium	0.013	PO
Bifidobacterium	0.02	NPO
Lactococcus	0.025	NPO
Candidate division TM7 uncultured bacterium	0.027	PO
Lactobacillus	0.03	NPO
Ruminococcaceae *Incertae Sedis*	0.046	PO

PI vs PO	**P value**	**Higher in**
Bifidobacterium	0.001	PI
Eremococcus	0.001	PI
Facklamia	0.001	PI
Alloiococcus	0.001	PI
Atopostipes	0.001	PI
Trichococcus	0.001	PI
Carnobacteriaceae uncultured	0.002	PI
DA101 soil group uncultured *Verrucomicrobia*	0.002	PO
Corynebacterium	0.003	PI
Exiguobacterium	0.003	PO
DA101 soil group uncultured bacterium	0.003	PO
Lactobacillus	0.004	PI

Ruminococcus	0.009	PI
Flavobacterium	0.009	PO
Massilia	0.012	PO
Sphingomonas	0.014	PO
Psychrobacter	0.019	PI
Prevotella	0.02	PI
Candidatus *Saccharimonas*	0.02	PI
Variovorax	0.02	PO
Tumebacillus	0.021	PO
Hymenobacter	0.022	PO
Dietzia	0.023	PI
Arthrobacter	0.028	PI
Clostridium sensu stricto 1	0.036	PI

NPI= No prep indoor; NPO=No prep outdoor; PI= Prep indoor; PO= Prep outdoor

Table S2: Corrected P values for Kruskal Wallis test on teat swab samples.

Teat		
NPI vs NPO	**P value**	**Higher in**
Trichococcus	0.012	NPI
Incertae Sedis	0.013	NPO
Acidimicrobiales uncultured bacterium	0.014	NPO
Peptostreptococcaceae uncultured	0.015	NPO
Marinospirillum	0.02	NPI
Proteiniphilum	0.021	NPI
Erysipelothrix	0.025	NPI
Exiguobacterium	0.025	NPO
Corynebacterium	0.039	NPI
Psychrobacter	0.039	NPI
Arenimonas	0.039	NPO
Betaproteobacteria uncultured bacterium	0.04	NPO
uncultured *Mycobacteriaceae* bacterium	0.042	NPO
Xanthomonadales uncultured bacterium	0.042	NPO
Eremococcus	0.044	NPI
Blastocatella	0.049	NPO
Verrucomicrobiaceae uncultured	0.049	NPO
Ferruginibacter	0.05	NPO
NPI vs PI	**P value**	**Higher in**
Variovorax	0.033	NPI
Bifidobacterium	0.04	PI
Acidimicrobiales uncultured bacterium	0.043	NPI
Devosia	0.043	NPI
NPO vs PO	**P value**	**Higher in**
No significant differences	NA	
PI vs PO	**P value**	**Higher in**
Proteiniphilum	0.001	PI
Jeotgalicoccus	0.001	PI
Eremococcus	0.001	PI
Facklamia	0.001	PI
Carnobacteriaceae uncultured	0.001	PI
Erysipelothrix	0.001	PI
Marinospirillum	0.001	PI
Blastocatella	0.001	PO
Acidimicrobiales uncultured.bacterium	0.001	PO
uncultured *Mycobacteriaceae* bacterium	0.001	PO
Dyadobacter	0.001	PO
Ferruginibacter	0.001	PO
Devosia	0.001	PO

Methylobacterium	0.001	PO
Variovorax	0.001	PO
Stenotrophomonas	0.001	PO
Xanthomonas	0.001	PO
DA101 soil group uncultured bacterium	0.001	PO
Corynebacterium	0.002	PI
Ruminococcus	0.002	PI
Psychrobacter	0.002	PI
uncultured Acidobacteria bacterium	0.002	PO
Clavibacter	0.002	PO
Gaiellales uncultured bacterium	0.002	PO
Chitinophagaceae uncultured	0.002	PO
Exiguobacterium	0.002	PO
Sphingomonas	0.002	PO
Betaproteobacteria SC.I.84 uncultured bacterium	0.002	PO
Arenimonas	0.002	PO
Xanthomonadales uncultured bacterium	0.002	PO
DA101 soil group uncultured *Verrucomicrobia* bacterium	0.002	PO
Verrucomicrobiaceae uncultured	0.002	PO
Bifidobacterium	0.003	PI
Spirochaetaceae uncultured	0.003	PI
Acidobacteria Subgroup 6 uncultured bacterium	0.003	PO
Hymenobacter	0.003	PO
Pedobacter	0.003	PO
WD2101 soil group uncultured bacterium	0.003	PO
Rhizobium	0.003	PO
Chthoniobacter	0.003	PO
Alloiococcus	0.004	PI
Trichococcus	0.004	PI
Rhodococcus	0.004	PO
Bradyrhizobium	0.004	PO
Prevotella	0.005	PI
Brevundimonas	0.005	PO
Massilia	0.005	PO
Spirosoma	0.006	PO
Nocardioides	0.007	PO
Treponema	0.009	PI
Kandleria	0.009	PO
uncultured *Parabacteroides* sp.	0.011	PI
Blautia	0.011	PI
Halomonas	0.015	PI
Cellvibrio	0.026	PO
Peptostreptococcaceae Incertae Sedis	0.03	PO
Lachnospiraceae uncultured	0.032	PI
Anaerotruncus	0.032	PI

Comamonadaceae uncultured	0.047	PO
Atopobium	0.05	PO

NPI= No prep indoor; NPO=No prep outdoor; PI= Prep indoor; PO= Prep outdoor

Table S3: Corrected P values for Kruskal Wallis test on faecal samples.

Faeces	P value	Higher in
Prevotella	0.001	Indoor
Bacteroides	0.002	Indoor
Saprospiraceae uncultured	0.002	Indoor
RF16 uncultured bacterium	0.005	Outdoor
Incertae Sedis	0.005	Outdoor
Ruminococcaceae uncultured	0.009	Outdoor
Lachnospiraceae uncultured	0.012	Indoor
Phascolarctobacterium	0.012	Indoor
Sutterella	0.012	Indoor
Ruminobacter	0.012	Indoor
Spirochaetaceae uncultured	0.012	Indoor
uncultured *Parabacteroides* sp.	0.021	Indoor
Ruminococcus	0.021	Indoor
Treponema	0.021	Indoor
RF9 uncultured bacterium	0.021	Indoor
Peptostreptococcaceae uncultured	0.021	Outdoor
Fibrobacter	0.027	Indoor
Incertae Sedis	0.027	Indoor
Phocaeicola	0.027	Outdoor
Alloprevotella	0.027	Outdoor
Paludibacter	0.036	Outdoor
Prevotellaceae uncultured	0.036	Outdoor
Blautia	0.046	Indoor

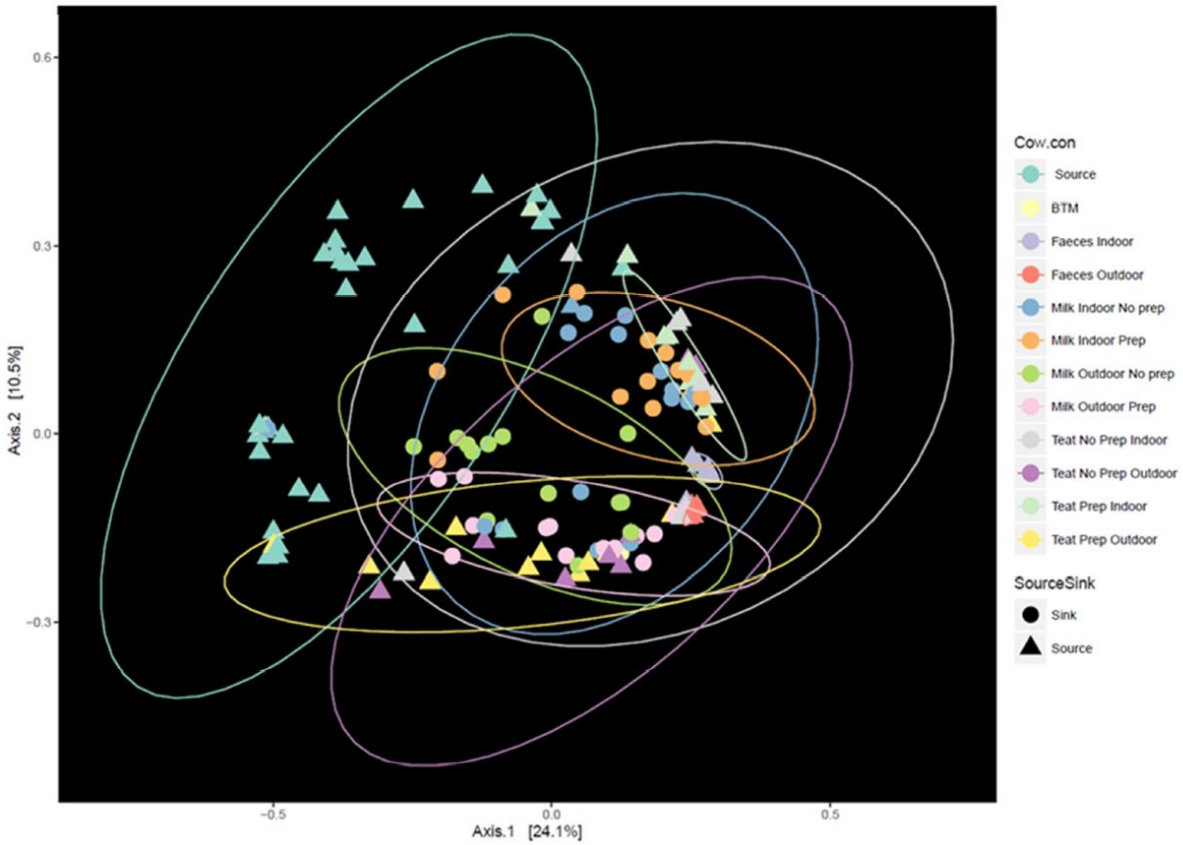


Fig.S1: Bray-Curtis PCoA of all samples coloured by farming practices and shaped based on source or sink. The cyan colour indicates environmental samples (soil, bedding, grass and silage).

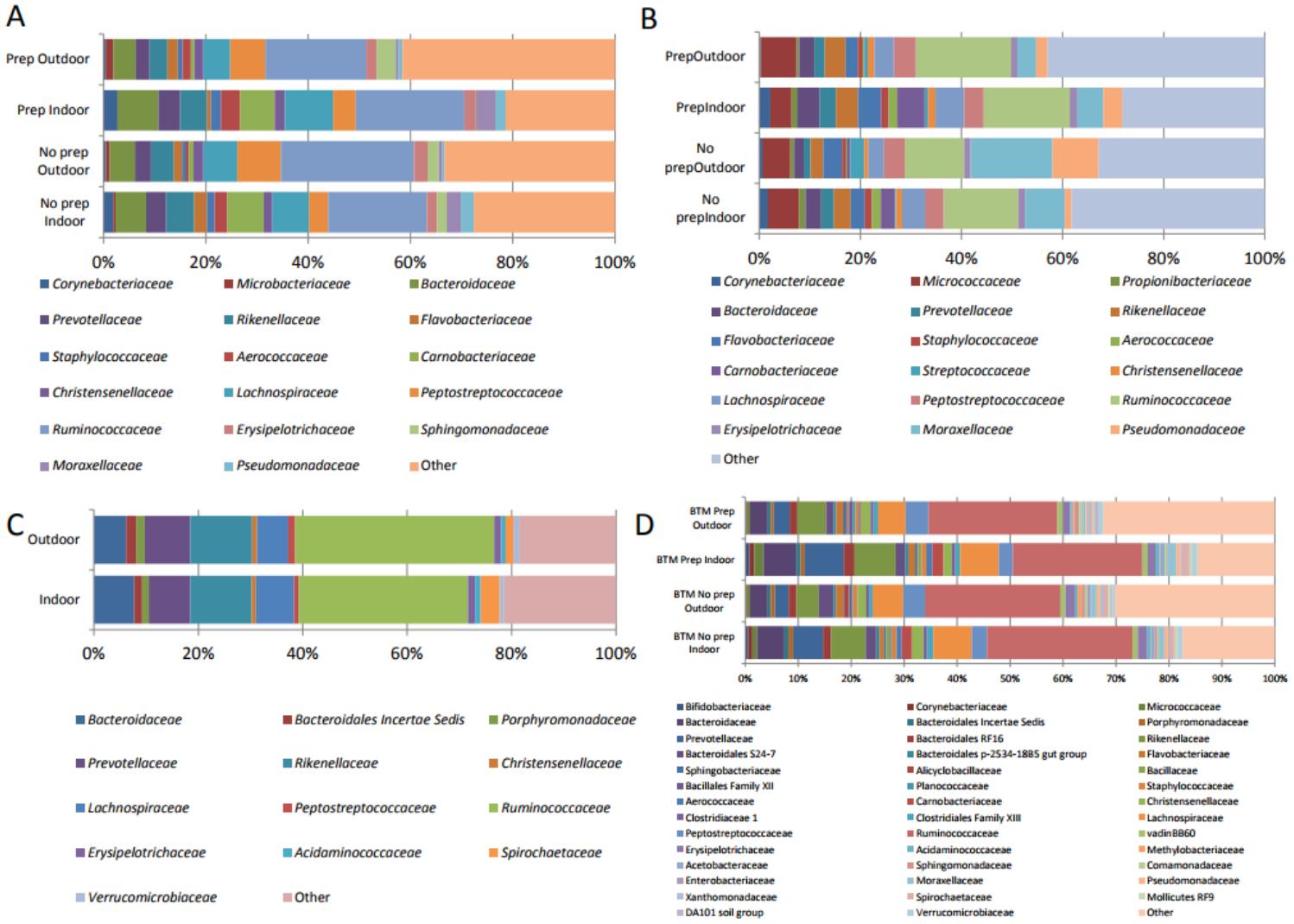


Fig.S2: Family present in (A) teat surface samples, (B) individual milk samples, (C) faecal samples and (D) BTM samples.

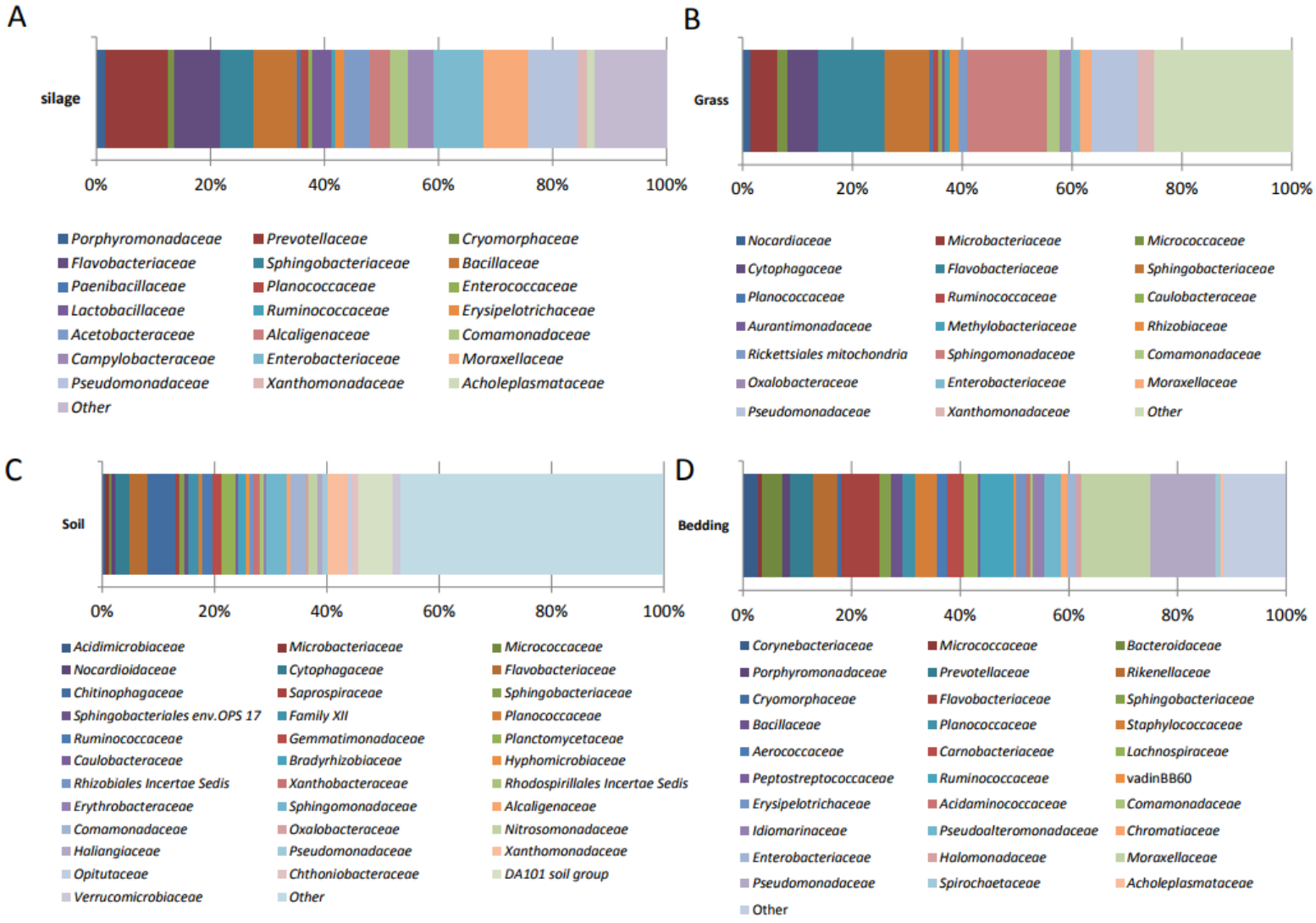


Fig.S3: Family present in environmental samples, (A) Silage, (B) Grass, (C) Soil and (D) Bedding samples.

Chapter 5

Metagenomic surveillance of the cheese production microbiome

5.0 Abstract

The microbial consortia present in food production environments can become part of the microbial composition of the final food products. Microorganisms which colonise food production facilities are thus an important consideration for food producers from both a spoilage and food safety perspective. Here, shotgun metagenomic sequencing was used to characterise the microbiome of a cheese production facility before and after the production of a continental style cheese. By adopting this approach, we were able to detect bacteria and phage present in substrates, production facility surfaces and in the cheese itself. Taxonomic analysis demonstrated that the production plant surfaces harboured lactic acid bacteria (LAB) prior to cheese production, but that the identity of the LAB present on these surfaces changed after the cheese production to be dominated by the specific strains used in continental cheese production. Pathogenic bacteria were detected in substrate samples on production plant surfaces but, importantly, not in cheese samples. Notably, brine and process water samples were found to harbour lactococcal phage, a major cause of starter culture failure in the cheese industry. These results demonstrate that shotgun metagenomic sequencing has the potential to become a valuable tool for monitoring the microbiology of food production facilities.

5.1 Introduction

Food processing environments can play a major role in determining the microbial composition of food products (Bokulich and Mills 2013, Bokulich, Lewis et al. 2016). It is evident that these environments contain many distinct microbial communities (Bokulich and Mills 2013, Bokulich, Bergsveinson et al. 2015), originating from a variety of different sources such as raw materials, air and production staff (Montel, Buchin et al. 2014). Once microbes are introduced into these facilities they may occupy specific niches and persist. Advances in high-throughput DNA sequencing (HTS) technologies have permitted ever greater insights into the microbiome of foods such as milk (Doyle, Gleeson et al. 2017, Doyle, Gleeson et al. 2017), dairy products (O'Sullivan, Cotter et al. 2015, Walsh, Crispie et al. 2017) and, indeed, food production facilities (Bokulich and Mills 2013, Bokulich, Bergsveinson et al. 2015, Hultman, Rahkila et al. 2015, Calasso, Ercolini et al. 2016). Although amplicon-based sequencing has been utilised primarily for this purpose thus far (Doyle, O'Toole et al. 2017), one of the issues with this approach is that, at best, it can only provide information regarding the bacterial (16S rRNA sequencing) or fungal (ITS sequencing) components of each environment. To simultaneously study the taxonomy of a microbiome in its entirety, including phage and viruses in these food related environments, as well as the functional potential thereof, shotgun metagenomic sequencing has been carried out on the microbiome of fermented food products (Walsh, Crispie et al. 2017) such as cheese (Quigley, O'Sullivan et al. 2016), kefir (Walsh, Crispie et al. 2016) and nunu (Walsh, Crispie et al. 2017). Although, this approach has not yet been applied to the dairy production environment, it has been used recently to characterise the microbiome present

in hospitals (Lax, Sangwan et al. 2017), in space craft assembly facilities (Bashir, Ahmed et al. 2016) and in beef production environments (Yang, Noyes et al. 2016).

In this proof of concept study, we aimed to characterise the cheese production facility microbiome, both before and after production of a Continental-style cheese. Additionally, the microbiome of the raw substrates used and the final cheese produced was also determined. To achieve this, we performed shotgun metagenomic sequencing of the microbiome. We found that the cheese production environment is dominated by starter bacteria retained from previous cheese production processes, with phage that target starter bacteria, and that have the potential to negatively impact on cheese production, also being present in high levels throughout this environment.

5.2 Materials and methods

5.2.1 Sample collection and extraction

Cheese production and food/environmental sampling was conducted in a pilot scale cheese production plant in Cork, Ireland. This production plant is used to produce cheeses of different varieties. For the purposes of this study, cheeses were produced based on a Swiss-type model as described previously (O'Sullivan, Cotter et al. 2015). With the starter cultures Steptocuccus thermophilus DPC6986 and Lactobacillus helveticus DPC6865, as well as Propionibacterium freudenreichii DPC6451. To produce the cheese, raw milk was collected from a local dairy farm and standardised to a protein-to-fat ratio of 1.01:1. Swab samples were collected from the cheese production facility before and after cheese production.

Swabs were dipped in a solution of 3 mL NaCl (0.09%) prior to swabbing, to improve recovery (O'Sullivan, Cotter et al. 2015). The production surfaces that were sampled before and after cheese production were as follows: milk vat, curd knife (knife) and draining table (drain). The cheese mold was sampled only before the curd was molded. A total of three swabs were used to swab each production facility surface. To ensure full swab head contact with the surfaces, overlapping "S" strokes of the swab with rotation of the swab were conducted with 4 rotations per swab. Three swabs from each surface were then pooled in a 9 mL NaCl solution in a sterile 15 mL falcon tube (Sarstedt, Ireland) and vortexed for 2 minutes. This resulted in one sample pool per surface sampled, both before and after production. These samples were then processed as described previously (Doyle, Gleeson et al. 2017).

In additon to the swab samples, process water from the facility, standardised milk and brine were also collected aseptically. Standardised milk (60 mL) was collected aseptically using 30 mL sterile blue dippa sample tubes (Ocon chemicals). For process water, the tap and handle of process water spouts were sterilised using 80% isopropyl alcohol wipes. Following this, 250 mL of process water was discharged into a sterile collection vessel. Brine samples (250 mL) were also collected using 30 mL sterile blue dippa sample tubes. DNA was extracted from the standardised milk as described previously (Doyle, Gleeson et al. 2017). For brine and process water extractions, 250 mL of each sample was filtered through a 0.45 µM nitrocellulose filter membrane (Merck Millipore). After filtration, the membranes were aseptically cut into microbead tubes (Powersoil kit, Mo Bio) using a sterile scissors (Medguard, Meath, Ireland) and forceps (VWR). Finally, cheese samples were sampled aseptically at day 60 post production, using a cheese trier. Following this, 5 g of cheese was

added to 50 mL of 2% trisodium citrate and this mixture was homogenised in a sterile stomacher bag. The resulting homogenate was aseptically transferred to 15 mL Falcon tubes (Sarstedt) and centrifuged at 5,444 × g for 30 min at 4°C. The supernatant was discarded and the cell pellet was washed with sterile PBS and transferred to a 2mL eppendorf tube where it was centrifuged at 14,000 x g for 1 minute. The resulting pellet was then dissolved in the lysis solution from the microbead tubes.

After pre-processing of samples, extractions were carried out using the Powersoil kit by adding it to the microbead tubes and using the standard kit protocol. DNA was quantified using the Qubit (Bio-Sciences, Dublin, Ireland), along with the high sensitivity DNA quantification assay kit (BioSciences).

5.2.2 DNA library preparation and sequencing

Extracted DNAs, with the exception of that from the three cheese samples, was subjected to multiple displacement amplification (MDA) to compensate for low starting concentrations. Sample preparation for this amplification process was carried out in a UV hood, using the Qiagen REPLI-g single cell kit (Qiagen, Manchester, England) as described previously (Bashir, Ahmed et al. 2016). All tubes and equipment were UV sterilised prior to the MDA. Amplifictions were conducted as per kit manual, using 1 µL of input metagenomic DNA. Three cheese samples were processed for sequencing without MDA . DNA was cleaned up using Agencourt AMPure XP magnetic beads. DNA was then quantified and diluted for library preparation, which was carried out using the Nextera XT DNA Library Preparation Kit

(Illumina). Samples were then sequenced on the Illumina NextSeq sequencing platform in the Teagasc sequencing facility, using a 2x250bp cycle V3 kit, following standard Illumina sequencing protocols.

5.2.3 Bioinformatic analysis

Raw reads from whole-metagenome shotgun sequencing were filtered on the basis of quality and quantity and trimmed to 200 bp with a combination of Picardtools https://github.com/broadinstitute/picard) and SAMtools. The Kraken classifier (Wood and Salzberg 2014) was used to determine taxonomic differences in the microbial composition of the cheese production facility and associate samples. SUPER-FOCUS (Silva, Green et al. 2015) was used to analyse the functional profiles of the microbial communities in this surveillance. PanPhlAn analysis (Scholz, Ward et al. 2016) was used to characterise the strains of starter bacteria present in the production facility, the substrates or the products.

The presence of Mycobacterium tuberculosis specific genes was detected using Bowtie2 (Langmead and Salzberg 2012). In addition potential antimicrobial resistance genes and virulence factors were identified by aligning reads from each sample against MEGARes (Lakin, Dean et al. 2017), Microbial Virulence Database MvirDB (Zhou, Smith et al. 2006) databases respectively using Bowtie2 (Langmead and Salzberg 2012).

5.3 Results

5.3.1 Shotgun sequencing of the cheese production environment microbiome

A total of 40 cheese and cheese production environmental samples were prepared for shotgun metagenomic sequencing. This included samples from the production facility surfaces (n=21; milk vat before=3, milk vat after=3, curd knife before=3, curd knife after=3, draining table before=3, draining table after=3, mold=3), samples from milk or water used in the manufacturing process (n=9; brine=3, milk =3 and process water=3), and cheese samples to represent the final product (n=6; cheese whole genome amplified (WGA)=3, cheese non-WGA=3) and a kit extraction MDA amplified negative control (n=1). There was an average of $32.7x10^6$ reads per sample pre-filtering, with reads ranging from 20.0 to 46.6 $x10^6$ per samples (Fig.S1).

5.3.2 Alpha and beta diversity of production plant surfaces are altered following cheese-production

Alpha diversity of the microbial populations was calculated using the species level data from Kraken and the Shannon diversity index (Fig.1). Of the liquid samples, the process water microbiota had a higher alpha diversity than that of both the brine and milk samples. A pattern was observed whereby the microbiota of the various production surfaces had a higher Shannon diversity index after production compared to the pre-production index. The mould microbiota had a higher alpha diversity than all milk vat and draining table samples, but values were lower than that of curd knife samples. The microbiota of whey was found to have a higher alpha diversity than that of all cheese samples. Overall, the microbiota of the

curd knife after production had the highest alpha diversity and the cheese WGA sample displayed the lowest microbial diversity.

Beta diversity analysis highlighted differences in relatedness between the production plant microbiome surfaces and the cheese samples as well as the brine and water substrate samples (Fig. 2). The production facility samples displayed the largest variability in microbiome composition, as evidenced by a larger polygon surface area (Fig. 2). This variability decreased after production. Although milk samples cluster with the production facility samples, brine and water samples are clearly distinct. Cheese samples cluster away from production facility and substrate samples, with WGA cheese samples displaying more variability than non-WGA cheese samples. The negative control does not cluster with any of the cheese or cheese production environment samples.

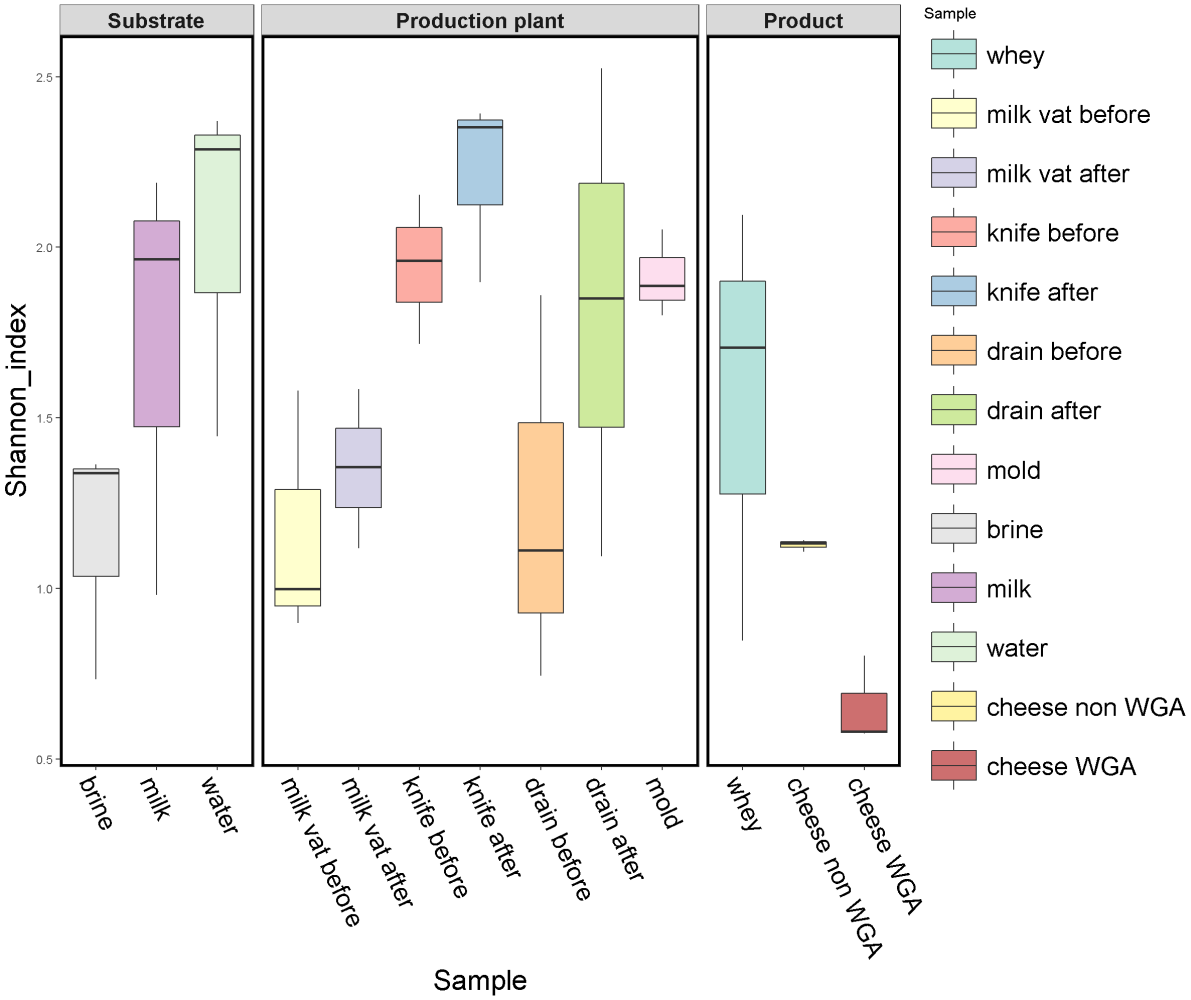


Fig.1: Alpha diversity boxplot depicting Shannon diversity index values for indicated sample types.

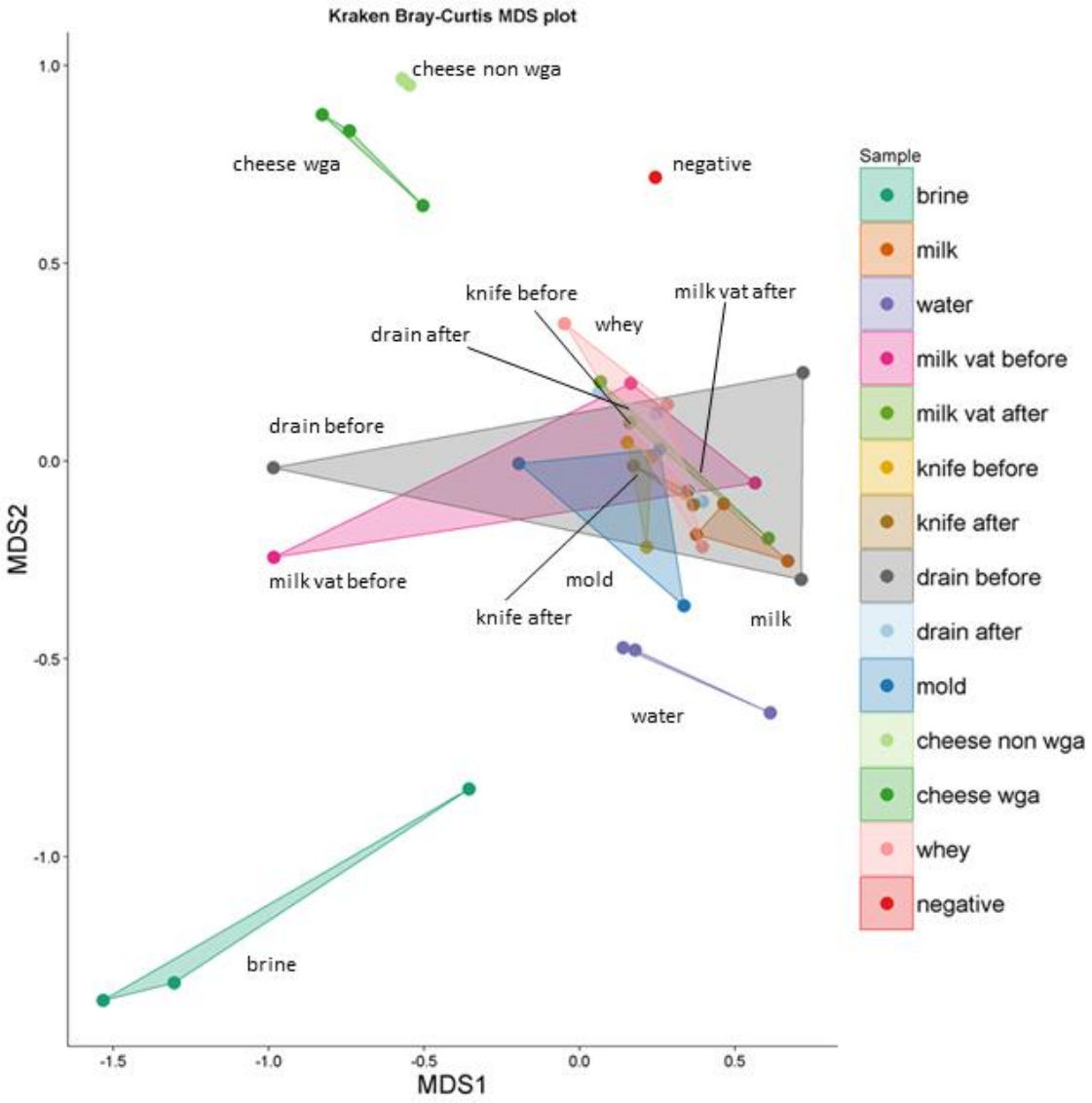


Fig.2: Bray-Curtis beta diversity MDS plot of Kraken species level analysis.

5.3.3 Species level compositional analysis of cheese production facility and associated metagenomes highlights the presence of resident lactic acid bacteria on production plant surfaces

Kraken analysis of sequence data facilitated the detection of bacterial and viral/(pro)phage DNA in the samples. After filtering and Kraken classification, an average of $1.96x10^6$ reads were assigned to species level for each environmental sample (Fig.S2). The relative abundance of each species in a sample type was calculated as a proportion of the mean number of reads per sample group. All species that were not present at an abundance above 1% in at least one sample group were classified as "other" (see Fig.3). All metagenomes contained species known to be of importance to dairy producers, including important starter/adjunct bacteria, potentially pathogenic bacteria, sporeformers and phage (Fig.4). In the production plant samples, the technologically important starter species *Lactococcus lactis* was present on production surfaces before cheese production had begun, but proportions were reduced post production. *L. lactis* was not used a starter for the Continental-type cheese produced and its presence presumably represents a carry-over from previous cheese manufacturing runs. These surfaces also contained *L. helveticus* and *S. thermophilus*, which were employed in the Continental-type cheese production run, in high proportions both before and after cheese production (Fig.4(A)). Both of these species were also found in process water. Cheese samples were dominated by *S. thermophilus*, and *L. helveticus* as well as the third species employed in the cheese manufacturing process, *Propionibacterium freudenreichii*, regardless of the manner in which samples were prepared (Fig.4(A)).

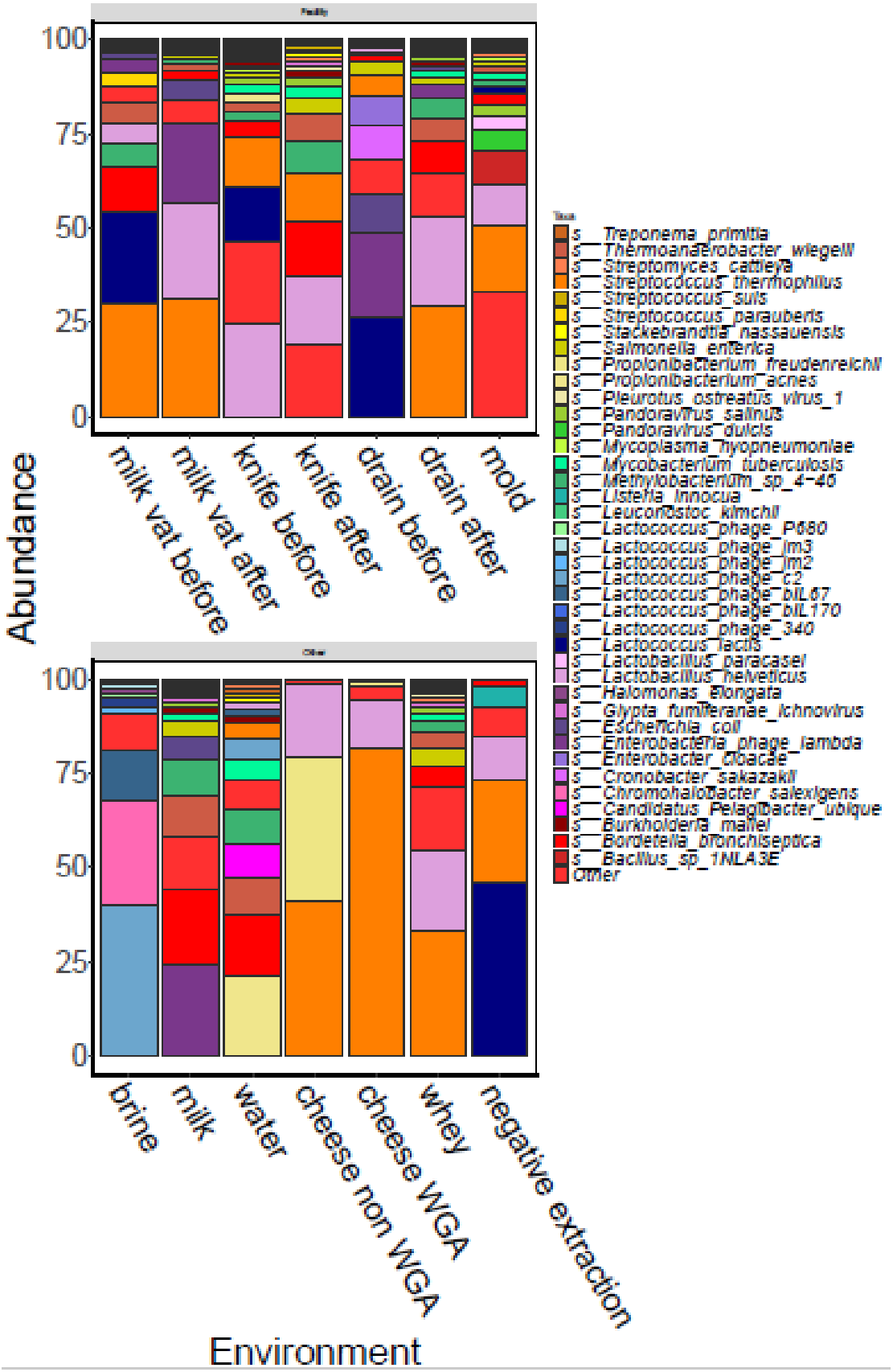


Fig.3: Kraken species level taxonomic composition of (A) the production facility surfaces and (B) substrate and products.

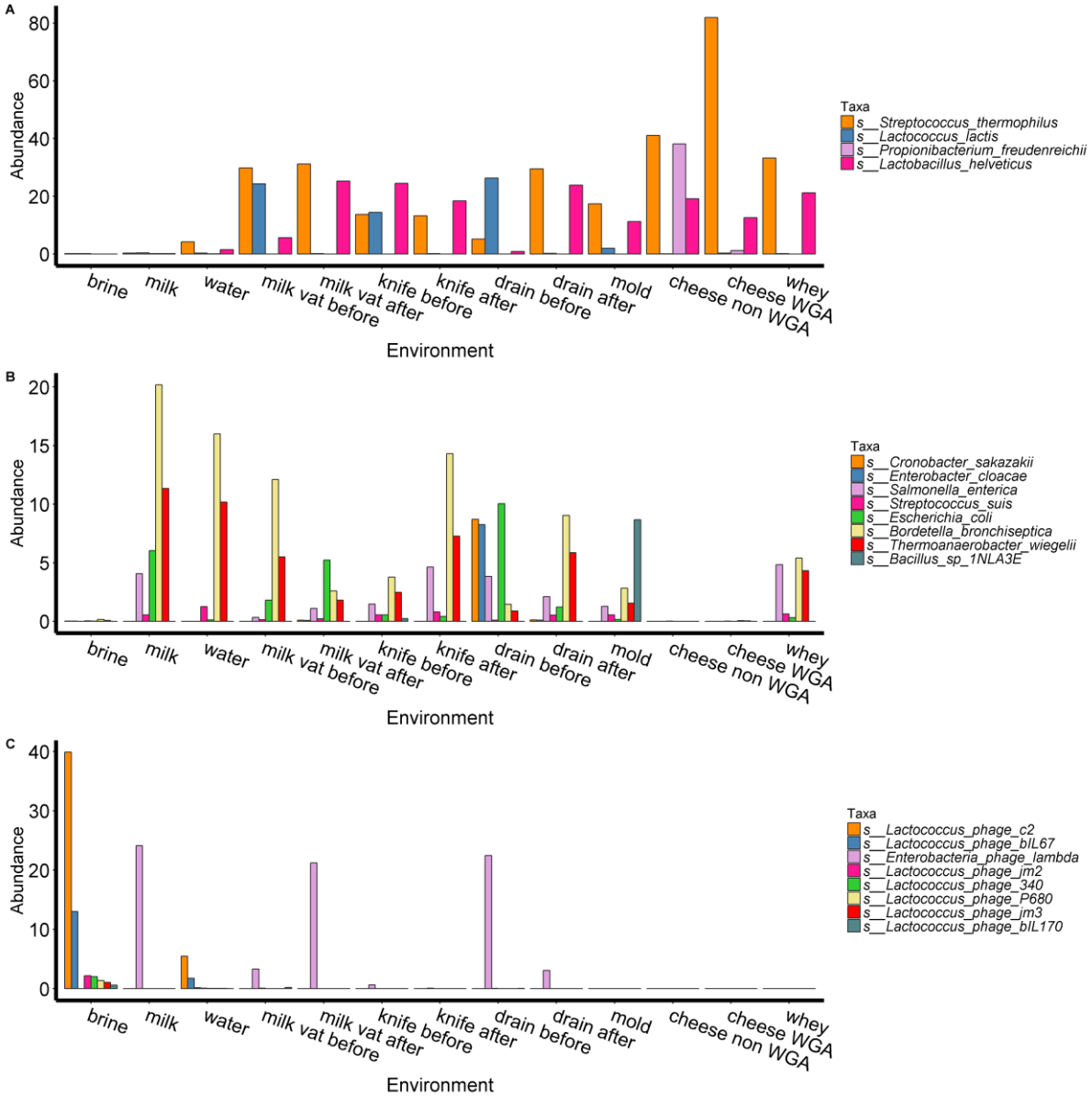


Fig.4: Bar plots depicting the differential abundance of (A) technologically important bacteria, (B) pathogenic and spoilage bacteria and (C) phage in all sample environments.

The distribution of well characterised pathogenic and potentially spoilage associated species was also investigated (Fig.4 (B)). Of the pathogenic species detected, reads assigned to *Salmonella enterica* and *Mycobacterium tuberculosis* were the most prevalent across all samples. Although the reads assigned to these species were detected in all samples with the exception of brine and cheese, further investigation highlighted that the putative *M. tuberculosis* reads had not been correctly assigned. More specifically, Bowtie2 failed to detect the presence of homologs of *hsp* 65, *DnaK* and *DnaJ* from *M. tuberculosis* H37Rv and *M. tuberculosis* was also not detected when using Metaphlan of the PanPhlAn analysis. *Cronobacter sakazakii* and *Enterobacter cloacae* were both detected above 5% in draining table samples before production while *Streptococcus suis* was also detected in substrate samples, production plant surfaces and in whey. Furthermore, potentially pathogenic species such as *Bordetella bronchiseptica* were detected in substrate samples (milk and process water), on production surfaces and in whey samples. Notably, no pathogenic species were detected in cheese samples. With regard to potential spoilage bacteria, the thermophilic sporeformer *Thermoanaerobacter wiegelii* was detected throughout the production plant, in milk and process water samples and in whey.

With regard to the distribution of (pro)phage, brine samples were found to contain a considerable proportion of phage sequences with homology to the *Lactococcus* phages C2, bIL67, bIL170, jm2 and jm3 (Fig. 4 (C)). *Lactococcus* phage c2 and biL67 were also detected in process water samples. These are likely to represent phage rather than prophage sequences as these samples did not contain high proportions of *Lactococcus*. The standardised milk samples contained high proportions of *Enterobacteria* phage. *Enterobacteria* phage sequences were detected in milk vat and draining table samples, and

were present at above 20% relative abundance in milk vat samples after production and draining table samples before production. The presence of these phage coincided with samples in which *E. coli* assigned reads were highest. Cheese samples were not found to contain any phage.

The negative control had a much lower number of reads classified in comparison to the other WGA metagenomes and indeed the non-WGA cheese sample. With only 52,641 reads were classified to species level in the corresponding negative control (Fig.S2). The negative control was found to have reads that were assigned to starter bacteria L. helveticus and S. thermophilius as well as L. lactis. In addition, it contained reads that were assigned to Listeria innocua.

5.3.4 Strain-level analysis to characterise starter bacteria in samples

The shotgun metagenomic data was also employed in conjunction with PanPhlAn for a strain level analysis of *L. helveticus* and *S. thermophilus* populations in each sample (Fig.S3). PanPhlAn detected *L. helveticus* in 17 out of 39 samples. The cheese strains group closely with *L. helveticus* GCF000015385 (Fig.S3 (A)). Indeed, among the other samples, all but one loosely cluster with the *L. helveticus* GCF000015385 genome. The outlier originated from a draining table sample prior to production. PanPhlAn detected *S. thermophilus* in 18 out of the 39 samples. The cheese sample strains group closely with *S. thermophilus* GCF000698885 strain, with different strains detected in brine and a curd knife sample collected after production.

5.3.5 Detection of virulence factors and antibiotic resistance highlights the distribution of virulence factors in the cheese production microbiome

Short read alignment was used to detect the presence of genes associated with microbial virulence factors and with antibiotic resistance. The highest number of antibiotic resistance genes was found in the cheese samples (Fig.S4 (A)). The brine, whey, draining table and milk vat samples also had a high number of reads of these genes, whereas process water and milk samples had a mean read count of less than 100 for antibiotic resistance genes. Rifampin resistance had the highest number of reads attributed to it, with the majority of these coming from brine and cheese samples. A list of the antibiotic resistance genes detected in each metagenome can be found in (Table.S1). Microbial virulence factors were found in high abundance throughout the cheese production plant and associated environments (Fig.S4 (B)). The draining table before production contained the highest number of virulence factor encoding genes while water samples had the lowest abundance.

5.3.6 Functional analysis highlights the potential influence cheese production has on associated microbial metabolism

SUPER-FOCUS analysis was carried out to determine the functional profiles of the metagenomes present (Fig.5). A number of pathways were found to differ in abundance (Fig. 6). Pathways related to phage appear to be most abundant in brine samples, but are also present in high abundance on production plant surfaces (Fig.6). Lower levels of these phage associated pathways were observed in water, mould, whey and cheese samples. In the water sample, respiration and cell wall-associated pathways were elevated compared to other samples. Genes related to pathways for protein metabolism, a key feature of cheese

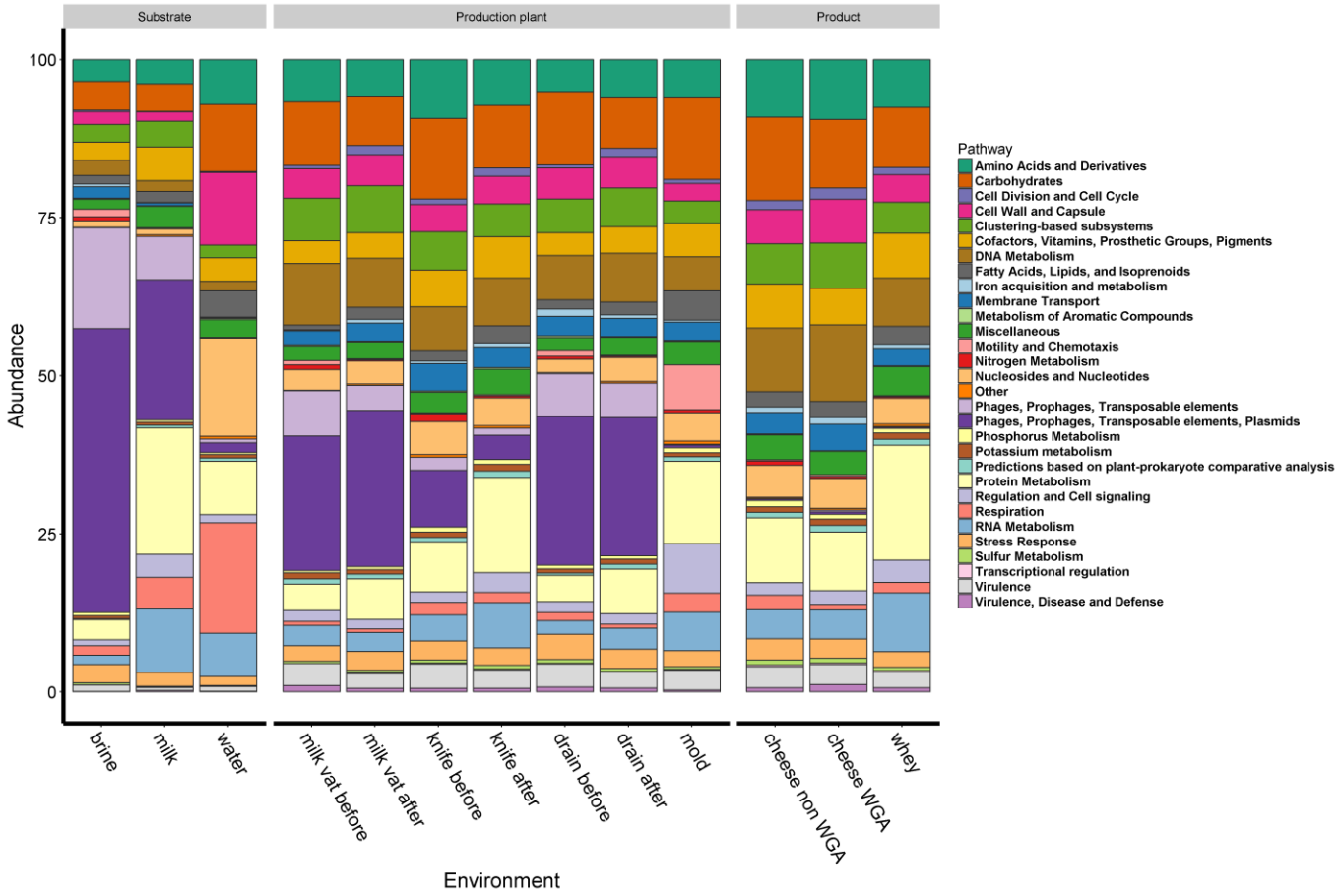


Fig. 5: SUPER-FOCUS pathway composition results.

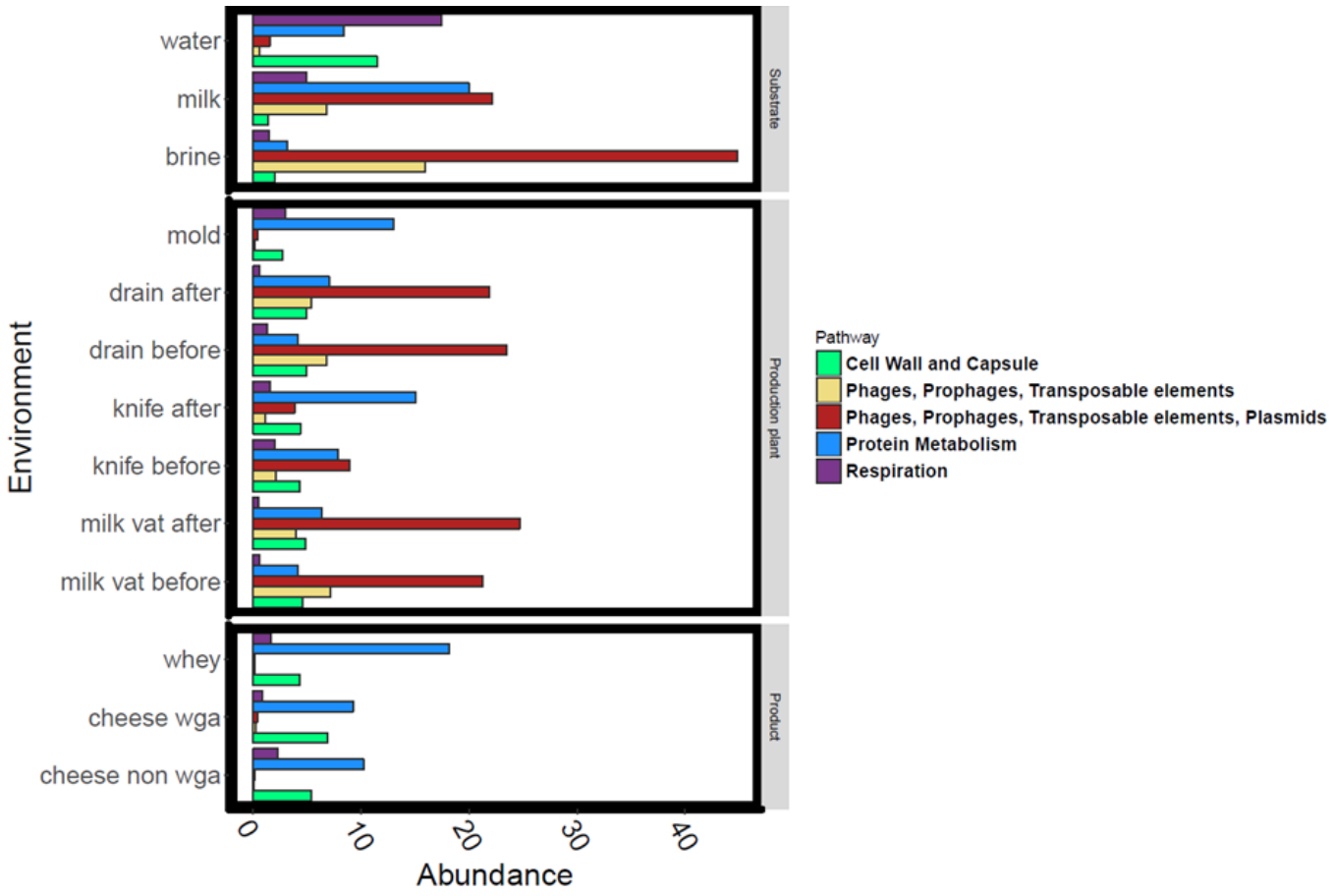


Fig.6: Differential abundance of SUPER-FOCUS pathways in sample groups.

production, were observed to be in greater abundance in milk, whey and both cheese samples. Additionally, protein metabolism pathways were found to have increased abundance on production plant surfaces after cheese production relative to before the commencement of production.

5.4 Discussion

High-throughput metagenomic sequencing has the potential to detect the presence of bacteria and viruses in food production environments. Until now, this technology has not been harnessed to analyse the food production environment. The detection of viruses and not readily cultivable bacteria in food and in food production environments is important from a food safety perspective as, for example, the majority of incidences of foodborne illness in the United States are caused by unknown agents (Scallan, Hoekstra et al. 2011). Additionally, from an industry perspective, the presence of phage in the production environment may cause issues with starter bacteria and lead to production losses (Mahony and van Sinderen 2015). This study may be the first step toward the application of shotgun sequencing for assessing, and thereby managing, the microbial population in food production facilities.

The alpha diversity for the production surface metagenomes was low in this study relative to that observed to a recent study conducted of an Italian dairy plant microbiota (Calasso, Ercolini et al. 2016). This suggests that the pilot production plant environment that was the focus of the present study is more controlled than industry production plant environments.

The alpha diversity of the production plants increased once production had taken place. This was expected, as the production plant surfaces would only have contained “house” microbes that are resilient and resistant to the implemented cleaning practises in the production plant prior to the initiation of the cheese production process. Once production commences, these surfaces are inoculated with microbes from raw materials and by starter bacteria. The cheese WGA samples were found to have the lowest alpha diversity out of all the samples; this is possibly due to a combination of a low diversity coupled with amplification biases introduced into these metagenomes during the MDA treatment that was not the case for non-WGA samples. This is also observed in the beta diversity plot; the cheese non WGA samples cluster tightly together, while the cheese WGA samples form a more variable cluster. The beta diversity plot also shows that there is a similarity shared in the production plant samples. The microbiota of these samples had more variability prior to production and this decreased after production. This change is again likely due to inoculation with substrate and starter culture bacteria.

Shotgun metagenomic analysis allowed an examination of the microbiome of all samples for the presence of technologically important bacteria, spoilage bacteria, potentially pathogenic microorganisms and phage. Technologically important starter bacteria were detected on production plant surfaces before cheese production. This is in agreement with some previous 16S based studies on the cheese production microbiota (Bokulich and Mills 2013, Calasso, Ercolini et al. 2016) and reinforces the view that the production plant microbiome consists of resident starter-type bacteria. PanPhlAn detected *L. helveticus* strains in 43% of metagenomes, with the majority of the strains detected in these metagenomes corresponding to *L. helveticus* GCF000015385 (DPC 4571), a strain used for continental

cheese production (Callanan, Kaleta et al. 2008). One other strain was detected from the draining table. The distribution of *S. thermophilus* strains was also analysed, with strains detected in 46% of metagenomes. Most of these strains grouped most closely with *S. thermophilus* GCF000698885 (ASCC 1275), which is a common dairy associated bacteria (Wu, Tun et al. 2014). Strains from one brine sample and a curd knife sample taken after production were observed to contain different strains of *S. thermophiles*, suggesting that other strains are resident constituents of the production plant microbiome.

The distribution of spoilage and potentially pathogenic bacteria was also examined. Milk, production surfaces, whey and water samples were found to contain *T. wiegelii*. This bacterium is a thermophilic sporeformer and members of this genus have previously been found in cheese process wastewater (Azbar, Dokgöz et al. 2009). *T. wiegelii* is a member of the Order Clostridia; members of which have previously been found to cause gas defects in cheese (Doyle, Gleeson et al. 2015). Potentially pathogenic bacteria such as *E. coli, S. enterica*, *E. cloacae*, *M. tuberculosis, S. suis*, *C. sakazakii* and *B. bronchiseptica* were detected in substrate samples, production plant surfaces, and whey samples, but not in cheese samples. *E. coli* is frequently found in raw milk and has previously been isolated from milk samples used for cheese production (Nobili, Franconieri et al. 2016) and *S. enterica* has previously been isolated from pasteurised milk (Olsen, Ying et al. 2004)and whey powder samples [36]. Recently, *E. cloacae* has been detected in raw ewe's milk used for cheese production (Cardinali, Osimani et al. 2017). This is the first reported incidence of this microbe detected in a food production environment. *S. suis* is a zoonotic pathogen which can be transmitted to humans (Lun, Wang et al. 2007), and has recently been detected in Italian cheese using culture independent analysis (De Filippis, La Storia et al. 2014). *C.*

sakazakii, formerly *Enterobacter sakazakii*, has previously been observed in food production facilities and in cheese (Kandhai, Reij et al. 2004). This opportunistic pathogen has the potential to cause illness in infants who consume contaminated powdered infant formula (Bowen and Braden 2006). *B. bronchiseptica* was present in high proportions in milk, process water samples and across the various the production surfaces despite not having been found in the cheese production environment previously. *B. bronchiseptica* is a zoonotic pathogenic which has caused illness in immune-compromised individuals (Dworkin, Sullivan et al. 1999). Finally, while reads were also initially assigned as *M. tuberculosis*, upon further examination no *M. tuberculosis* specific genes were detected in any sample. Additionally, Metaphlan did not detect the presence of this species. This result highlights that caution needs to be exercised when conducting species level surveillance of shotgun metagenomic datasets and further improvements to classifiers are required to improve reliability.

Although distribution of bacteria in food production plants has been examined previously using amplicon based approaches, the availability of shotgun metagenomic data sets has enabled the simultaneous detection of phage in the cheese production microbiome in this present study Fig.S5(C). Brine samples had a particularly high proportion of phage DNA present, with *Lactococcus* phage C2-like sequences being most plentiful. This is consistent with previous observations highlighting the prevalence of this phage in dairy plants (Rousseau and Moineau 2009, Marcó, Moineau et al. 2012). The primary host for this phage is the common dairy starter bacteria *L. lactis* subsp. *lactis* and the presence of this phage in such high abundance in this environment and indeed, throughout the production facility, is likely due to its target host being used as a starter for other cheese production processes, such as Cheddar production. *Lactococcus* phage JM2 and JM3 were also detected in the

brine samples. Both phage, which have narrow specificity for *L. lactis* subsp. *cremoris,* have previously been detected at Irish facilities (Mahony, Kot et al. 2013). *Enterobacteria* phage was also detected in some environments, such as standardised milk, milk vat and draining table samples, with their presence coinciding with the presence of *E. coli*.

While the approach taken in this present study has the advantage of being able to detect the presence of bacteria and phage in the same samples, there is potential for further improvements. There may be some potential biases associated with the application of MDA-facilitated shotgun metagenomic surveillance, some of which have been highlighted recently (Thoendel, Jeraldo et al. 2017), and include preferential amplification of lower GC DNA strands, reduced amplification of lower abundance DNA strands and smaller DNA strands, similar biases are associated with the application of HTS to amplicon sequencing (Thoendel, Jeraldo et al. 2017). The primary example of bias in the current study related to the *S. thermophilus*: *P. freudenreichii* ratio in cheese samples, with *S. thermophilus* DNA appearing to be preferentially amplified DNA in the cheese WGA samples. It is also important to be aware that DNA extraction kit contamination can be an issue when preparing libraries for HTS (Salter, Cox et al. 2014). Thus the inclusion of a negative control is important to examine the extent to which this is a problem in individual studies. In this present study, the negative control sample had the lowest number of classified reads at species level using Kraken. This control did have some reads attributed to starter bacteria which could be explained by false index pair assignment to this sample [28]. Also present in this sample was *L. innocua* and *Cotesia congregate bracovirus* but these were not present in levels above 1% of the total reads in any other sample, suggesting that these are the result of reagent contamination. It is worth noting that *Methylobacterium* was detected in this

study in the milk and water samples, in whey samples and on production surfaces. While this genus has been identified as a DNA extraction kit contaminant previously (Salter, Cox et al. 2014), it was not detected in the negative kit control in this instance, suggesting that this microbe is not present as a result of kit contamination. Furthermore, this microbe has been detected previously in Italian cheese samples (Dolci, Barmaz et al. 2009), and in industrial milk samples (Bracke, Van Poucke et al. 2014).

Short read alignment also facilitated the detection of genes associated with microbial virulence factors and antibiotic resistance. Cheese samples had the highest number of reads corresponding to antibiotic resistance genes. Rifampin resistance was the most dominant resistance mechanism detected. However, it is apparent that one must be cautious when interpreting this type of data as some of these genes are simply housekeeping genes that have the potential to confer resistance as a consequence of acquiring single nucleotide mutations and, thus, their presence does not necessarily reflect resistance. Innate resistance to specific antibiotics is a common trait among dairy starter bacteria and has been described previously (Hummel, Hertel et al. 2007). Microbial virulence factors were found to be widely distributed throughout production plant microbiomes. Draining tables were found to harbour the highest number of virulence factors, which is not surprising given the abundance of pathogens detected in this niche. However, care needs to be taken in interpreting virulence gene density data as genes with homology to fitness-associated genes in pathogens do not contribute to pathogenicity in other strains (Hill 2012).

In agreement with the taxonomic results, brine samples were found to contain the highest proportion of reads attributed to phage functionality. Pathways for protein metabolism were highest in milk, whey and cheese samples; this can be attributed to the proteolysis

that occurs during the cheese-making processes and in cheese maturation. Furthermore, protein metabolism pathways are also elevated on production plant surfaces after, relative to before, production.

5.5 Conclusion

Here we describe, for the first time, the application of high-throughput shotgun metagenomic sequencing to characterise the cheese production environment. We examine how the microbiome of the production plant surfaces changes before and after cheese production. In addition, we show the distribution of phage through this environment. These results highlight how HTS-based technologies could be applied to detect the presence of spoilage and pathogenic agents of viral or bacterial nature in food production environments. With further methodological developments and reduced cost and increased speed sequencing, there is potential for this technology to be widely applied by the food industry.

5.6 References

Azbar, N., et al. (2009). "Continuous fermentative hydrogen production from cheese whey wastewater under thermophilic anaerobic conditions." International Journal of Hydrogen Energy **34**(17): 7441-7447.

Bashir, M., et al. (2016). "Functional Metagenomics of Spacecraft Assembly Cleanrooms: Presence of Virulence Factors Associated with Human Pathogens." Frontiers in Microbiology **7**.

Bokulich, N. A., et al. (2015). "Mapping microbial ecosystems and spoilage-gene flow in breweries highlights patterns of contamination and resistance." eLife **4**: e04634.

Bokulich, N. A., et al. (2016). "A new perspective on microbial landscapes within food production." Current opinion in biotechnology **37**: 182-189.

Bokulich, N. A. and D. A. Mills (2013). "Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants." Applied and environmental microbiology **79**(17): 5214-5223.

Bowen, A. B. and C. R. Braden (2006). "Invasive Enterobacter sakazakii disease in infants." Emerg Infect Dis **12**(8): 1185.

Bracke, N., et al. (2014). "Identification of a microscopically selected microorganism in milk samples." Journal of Dairy Science **97**(2): 609-615.

Calasso, M., et al. (2016). "Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant." Food microbiology **54**: 115-126.

Callanan, M., et al. (2008). "Genome sequence of Lactobacillus helveticus, an organism distinguished by selective gene loss and insertion sequence element expansion." J Bacteriol **190**(2): 727-735.

Cardinali, F., et al. (2017). "Impact of thistle rennet from Carlina acanthifolia All. subsp. acanthifolia on bacterial diversity and dynamics of a specialty Italian raw ewes' milk cheese." Int J Food Microbiol.

De Filippis, F., et al. (2014). "A selected core microbiome drives the early stages of three popular Italian cheese manufactures." PloS one **9**(2): e89680.

Dolci, P., et al. (2009). "Maturing dynamics of surface microflora in Fontina PDO cheese studied by culture-dependent and-independent methods." Journal of applied microbiology **106**(1): 278-287.

Doyle, C. J., et al. (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." Int J Food Microbiol **197**: 77-87.

Doyle, C. J., et al. (2017). "High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions." Int J Food Microbiol.

Doyle, C. J., et al. (2017). "Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study." Applied and environmental microbiology **83**(2): e02694-02616.

Doyle, C. J., et al. (2017). "Metagenome-based surveillance and diagnostic approaches to studying the microbial ecology of food production and processing environments." Environmental microbiology.

Dworkin, M. S., et al. (1999). "Bordetella bronchiseptica infection in human immunodeficiency virus-infected patients." Clinical Infectious Diseases **28**(5): 1095-1099.

Hill, C. (2012). "Virulence or niche factors: what's in a name?" J Bacteriol **194**(21): 5725-5727.

Hultman, J., et al. (2015). "Meat processing plant microbiome and contamination patterns of cold-tolerant bacteria causing food safety and spoilage risks in the manufacture of vacuum-packaged cooked sausages." Applied and environmental microbiology **81**(20): 7088-7097.

Hummel, A. S., et al. (2007). "Antibiotic resistances of starter and probiotic strains of lactic acid bacteria." Applied and environmental microbiology **73**(3): 730-739.

Kandhai, M. C., et al. (2004). "Occurrence of Enterobacter sakazakii in food production environments and households." The Lancet **363**(9402): 39-40.

Lakin, S. M., et al. (2017). "MEGARes: an antimicrobial resistance database for high throughput sequencing." Nucleic acids research **45**(D1): D574-D580.

Langmead, B. and S. L. Salzberg (2012). "Fast gapped-read alignment with Bowtie 2." Nature methods **9**(4): 357-359.

Lax, S., et al. (2017). "Bacterial colonization and succession in a newly opened hospital." Science Translational Medicine **9**(391): eaah6500.

Lun, Z.-R., et al. (2007). "Streptococcus suis: an emerging zoonotic pathogen." The Lancet infectious diseases **7**(3): 201-209.

Mahony, J., et al. (2013). "Investigation of the relationship between lactococcal host cell wall polysaccharide genotype and 936 phage receptor binding protein phylogeny." Applied and environmental microbiology **79**(14): 4385-4392.

Mahony, J. and D. van Sinderen (2015). "Novel strategies to prevent or exploit phages in fermentations, insights from phage–host interactions." Current opinion in biotechnology **32**: 8-13.

Marcó, M. B., et al. (2012). "Bacteriophages and dairy fermentations." Bacteriophage **2**(3): 149-158.

Montel, M.-C., et al. (2014). "Traditional cheeses: rich and diverse microbiota with associated benefits." Int J Food Microbiol **177**: 136-154.

Nobili, G., et al. (2016). "Isolation of Shiga toxin-producing Escherichia coli in raw milk and mozzarella cheese in southern Italy." Journal of Dairy Science **99**(10): 7877-7880.

O'Sullivan, D. J., et al. (2015). "Temporal and spatial differences in microbial composition during the manufacture of a Continental-type cheese." Applied and environmental microbiology **81**(7): 2525-2533.

Olsen, S. J., et al. (2004). "Multidrug-resistant Salmonella Typhimurium infection from milk contaminated after pasteurization." Emerg Infect Dis **10**(5): 932.

Quigley, L., et al. (2016). "Thermus and the Pink Discoloration Defect in Cheese." mSystems **1**(3): e00023-00016.

Rousseau, G. M. and S. Moineau (2009). "Evolution of Lactococcus lactis phages within a cheese factory." Applied and environmental microbiology **75**(16): 5336-5344.

Salter, S. J., et al. (2014). "Reagent and laboratory contamination can critically impact sequence-based microbiome analyses." BMC biology **12**(1): 87.

Scallan, E., et al. (2011). "Foodborne illness acquired in the United States—major pathogens." Emerg Infect Dis **17**(1).

Scholz, M., et al. (2016). "Strain-level microbial epidemiology and population genomics from shotgun metagenomics." Nature methods **13**(5): 435-438.

Silva, G. G. Z., et al. (2015). "SUPER-FOCUS: a tool for agile functional analysis of shotgun metagenomic data." Bioinformatics **32**(3): 354-361.

Thoendel, M., et al. (2017). "Impact of Contaminating DNA in Whole-Genome Amplification Kits Used for Metagenomic Shotgun Sequencing for Infection Diagnosis." Journal of clinical microbiology **55**(6): 1789-1801.

Walsh, A. M., et al. (2017). "Translating Omics to Food Microbiology." Annual Review of Food Science and Technology **8**: 113-134.

Walsh, A. M., et al. (2017). "Strain-level metagenomic analysis of the fermented dairy beverage nunu highlights potential food safety risks." Applied and environmental microbiology: AEM. 01144-01117.

Walsh, A. M., et al. (2016). "Microbial Succession and Flavor Production in the Fermented Dairy Beverage Kefir." mSystems **1**(5): e00052-00016.

Wood, D. E. and S. L. Salzberg (2014). "Kraken: ultrafast metagenomic sequence classification using exact alignments." Genome biology **15**(3): R46.

Wu, Q., et al. (2014). "Genomic insights into high exopolysaccharide-producing dairy starter bacterium Streptococcus thermophilus ASCC 1275." Scientific reports **4**.

Yang, X., et al. (2016). "Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within the microbiome of the beef production chain." Applied and environmental microbiology **82**(8): 2433-2443.

Zhou, C., et al. (2006). "MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications." Nucleic acids research **35**(suppl_1): D391-D394.

Supplementary material

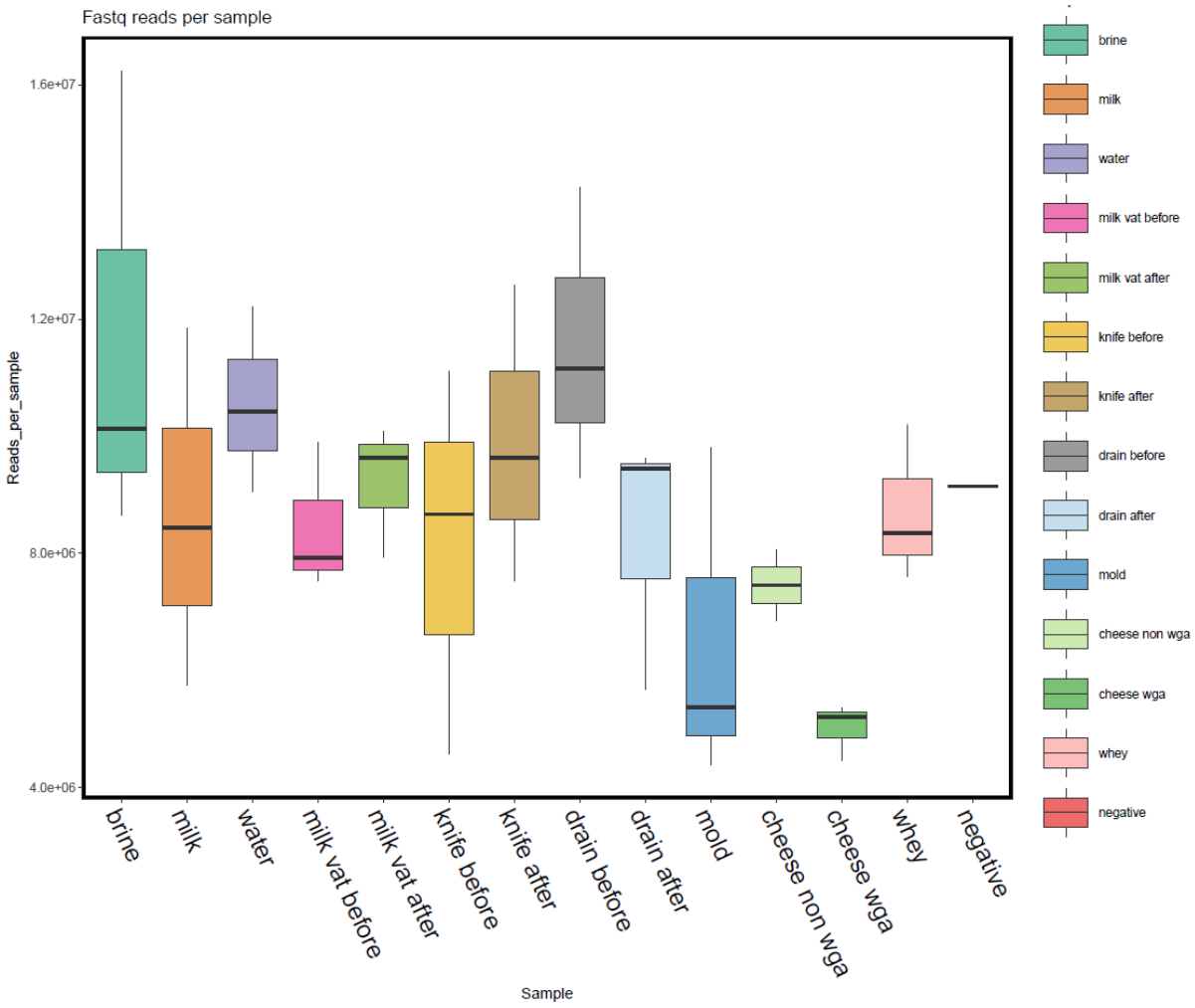


Fig.S1: Number of raw Fastq reads per sample group reflects equimolar pooling of samples.

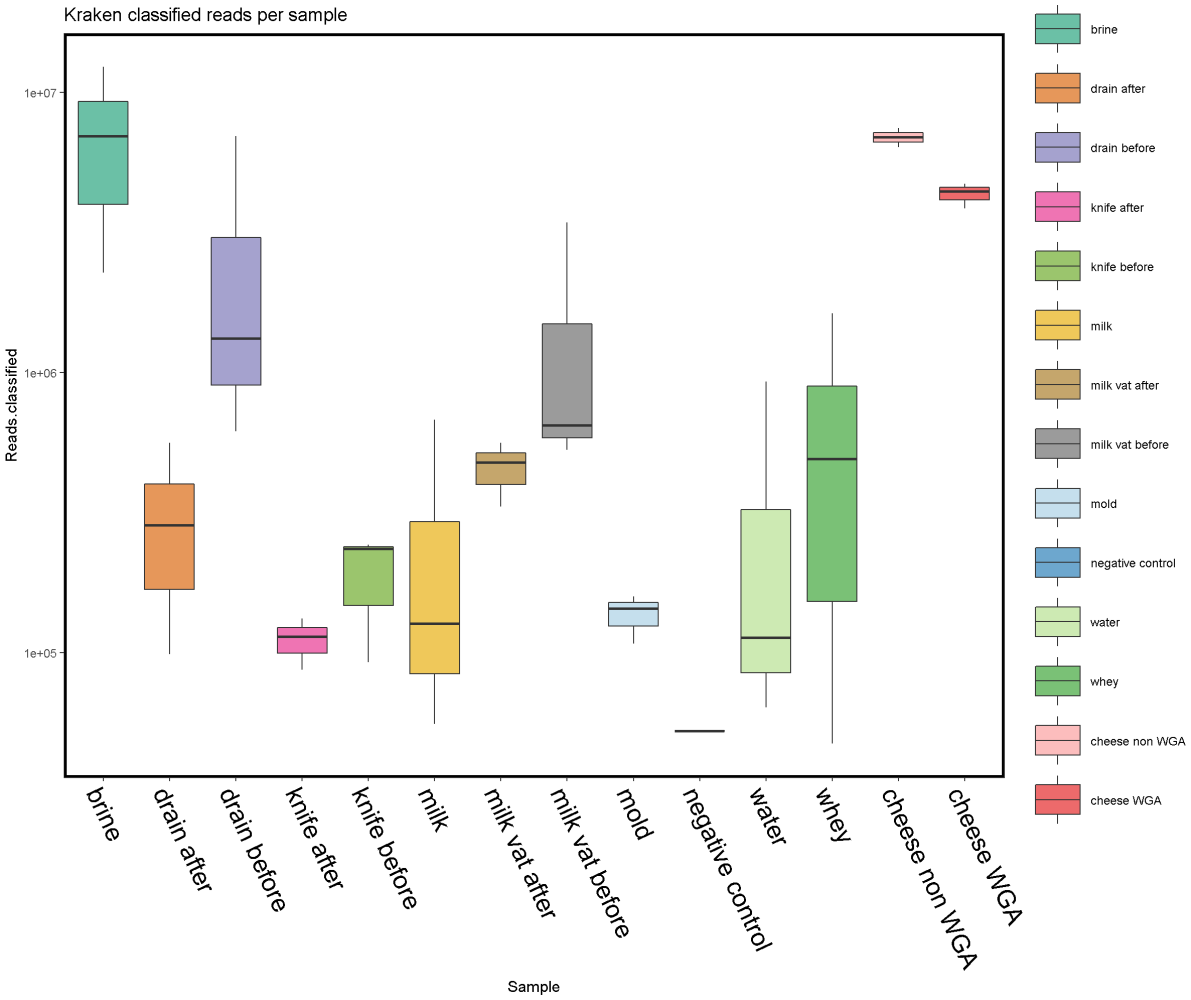


Fig.S2: Boxplot depicting the number of reads classified by Kraken in each metagenome.

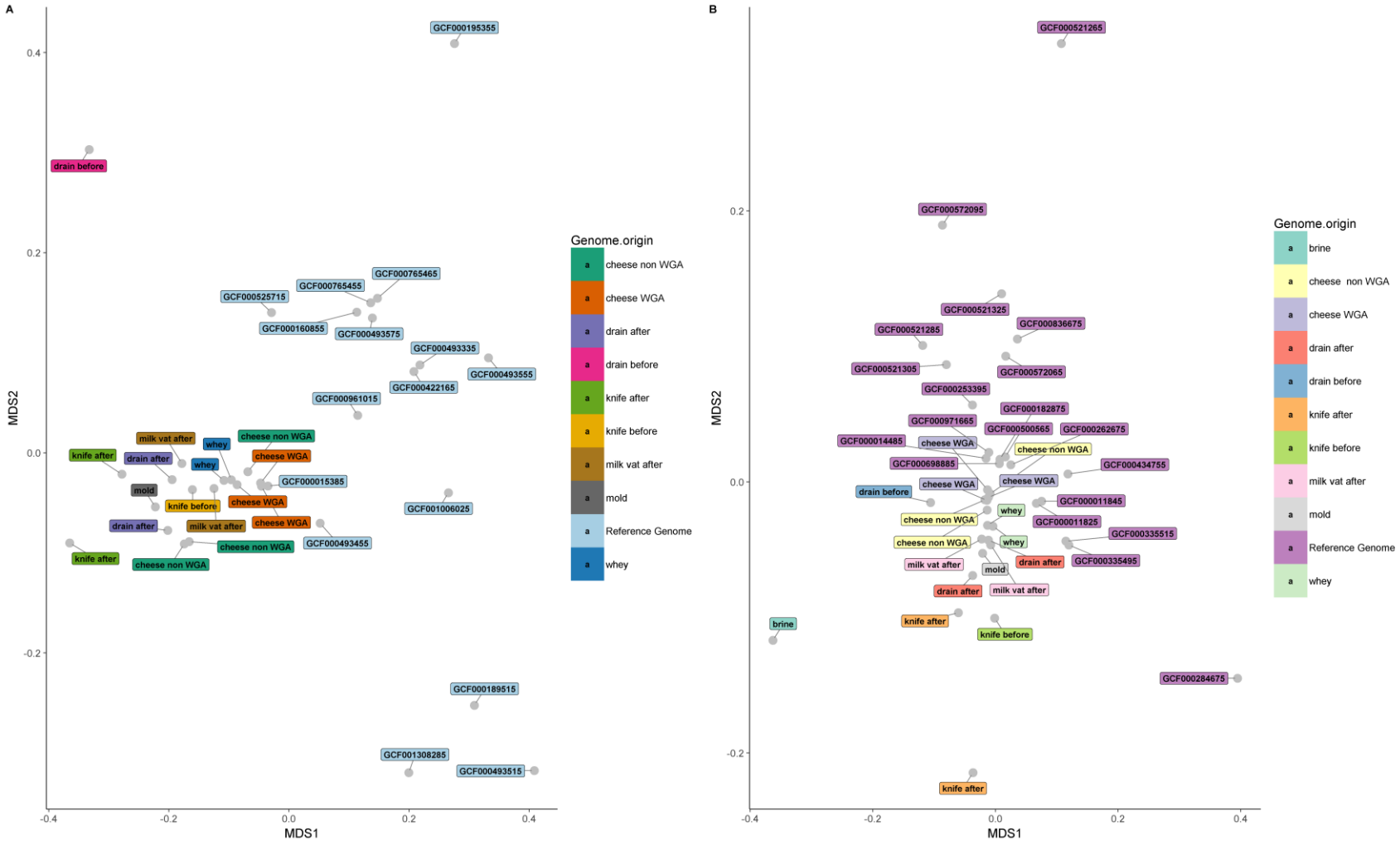


Fig. S3: MDS plot depiction the genome similarity shared between reference genomes and genomes detected in samples from this surveillance (A) *L. helveticus* and (B) *S. thermophiles.*

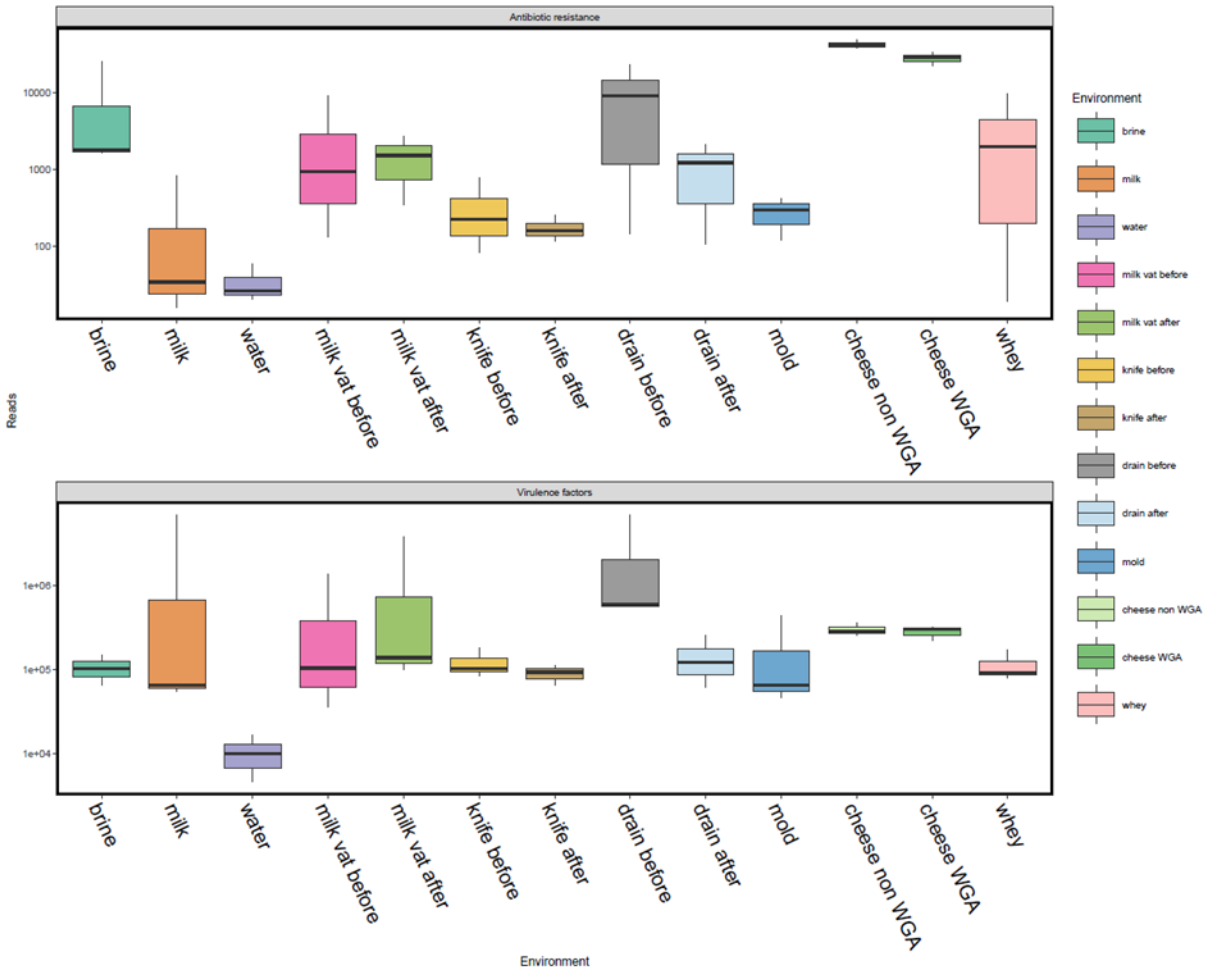


Fig.S3: Boxplots depicting the number of reads attributed to (A) antibiotic resistance genes and (B) microbial virulence factors.

Table S3: Reads attributed to antibiotic resistance genes

	brine	cheese non WGA	cheese WGA	drain after	drain before	knife after	knife before	milk	milk vat after	milk vat before	mold	water	whey
Rif|NC_002516.2.881699|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	311	295	166	0	0	0	2	0	1	0	0	3	1
CARD|pvgb|AE014075|3901532-3902762|ARO:3003438|Escherichia|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	8	194	164	12	16	63	61	0	4	84	2	14	91
Rif|CP002695.1|gene18|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	23	229	131	0	0	0	1	1	0	0	0	0	2
Elf|NC_002516.2.881697|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	36	124	64	0	0	0	3	0	0	0	14	0	0
CARD|phgb|M57437|0-1647|ARO:3002827|tlrC|MLS|ABC_transporter|TLRC	10	121	60	0	0	0	1	0	0	0	0	0	1
CARD|phgb|NC_007779|2586250-2589364|ARO:3000491|acrD|Aminoglycosides|Aminoglycoside_efflux_pumps|ACRD	12	0	0	0	172	6	0	0	0	0	0	0	0
CARD|pvgb|NC_002516|4767810-4769004|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	48	77	53	0	1	0	1	0	0	0	0	0	0
Elf|NC_002516.2.881718|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	28	90	56	0	0	0	1	0	0	0	0	0	0
MLS|TlrC|NC_016113|803268-384890|1623|MLS|ABC_transporter|TLRC	12	53	26	0	1	0	1	0	0	0	0	0	0
Rif|NC_008702.1.4609796|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	40	31	17	0	0	0	0	0	1	0	0	1	1
CARD|phgb|X63451|0-1653|ARO:3002828|srmB|MLS|Spiramycin_efflux_pumps|SRMB	81	3	1	0	2	0	0	0	0	0	0	0	0
CARD|pvgb|NC_002516.2|3556426-3559198|ARO:3003684|Pseudomonas|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	55	17	10	0	0	0	0	0	0	0	0	1	1
Flq|NC_002516.2.882800|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	57	10	7	0	0	1	0	0	0	0	0	0	1
CARD|pvgb|AP009048|3760295-3762710|ARO:3003303|Escherichia|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRB|RequiresSNPConfirmation	8	1	0	0	52	3	1	1	1	7	0	0	0
CARD|pvgb|110645304|5576027-5577917|ARO:3003685|Pseudomonas|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|parE|RequiresSNPConfirmation	28	27	10	0	1	0	1	0	0	0	0	0	1
Flq|CP002695.1|gene1275|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	6	37	21	0	0	0	1	0	0	0	0	0	1
ACou|NC_002516.2.879897|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	26	24	9	0	0	0	1	0	0	0	0	0	1
CARD|phgb|AB219524.1|1176-4338|ARO:3003699|mexQ|Multi-drug_resistance|Multi-drug_efflux_pumps|MEXQ	21	3	1	0	5	0	0	0	0	0	13	0	0
Flq|CP000647.1|gene2640|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	37	0	0	0	2	0	0	0	0	0	0	0	1

CARD|phgb|AB219523.1|1175-4286|ARO:3003705|mexN|Multi-drug_resistance|Multi-drug_efflux_pumps|MEXN	8	0	0	0	0	0	0	0	0	0	27	0	0
CARD|pvgb|CP002695|3866610-3867801|ARO:3001312|elfamycin|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	8	7	8	0	2	0	0	0	0	0	7	0	1
CARD|pvgb|AB003428.1|151-2416|ARO:3003702|Pseudomonas|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	12	10	11	0	0	0	0	0	0	0	0	0	0
Flq|NC_002516.2.879741|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	12	11	8	0	0	0	0	0	0	0	0	0	1
Rif|NC_003112.2.902240|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	13	5	5	0	0	2	1	0	0	2	0	0	4
ACou|CP000647.1|gene3444|Aminocoumarins|Aminocoumarin-resistant_DNA_topoisomerases|PARE|RequiresSNPConfirmation	7	0	0	0	24	1	0	0	0	0	0	0	0
Elf|CP002695.1|gene3614|Elfamycins|EF-Tu_inhibition|TUFAB|RequiresSNPConfirmation	7	10	3	0	1	0	0	0	0	0	7	0	1
Rif|NC_003197.1.1255679|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	26	0	0	0	2	0	0	0	0	0	0	0	0
Flq|CP000647.1|gene3437|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	11	0	0	0	16	1	0	0	0	0	0	0	0
Rif|CP000647.1|gene4402|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	21	1	1	0	2	1	0	0	0	0	0	0	0
Rif|CP001138.1|gene4362|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	22	0	0	0	0	0	0	0	0	0	0	0	0
CARD|phgb|U00096|2155262-2158385|ARO:3000793|mdtB|Multi-drug_resistance|Multi-drug_efflux_pumps|MDTB	2	0	0	0	9	0	0	0	1	0	0	0	0
Rif|FN543093.2|gene314|Rifampin|Rifampin-resistant_beta-subunit_of_RNA_polymerase_RpoB|RPOB|RequiresSNPConfirmation	4	3	2	0	1	0	0	0	0	1	0	0	0
Flq|NC_002695.1.916822|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	8	0	0	0	1	0	0	0	0	0	0	0	0
CARD|phgb|M80346|0-1656|ARO:3002817|carA|MLS|Macrolide_resistance_efflux_pumps|CARA	7	0	0	0	0	0	0	0	0	0	0	0	0
CARD|phgb|AF173226|351-1041|ARO:3003066|smeR|Multi-drug_resistance|Multi-drug_efflux_pumps|sme	6	0	0	0	0	0	0	0	0	0	0	0	0
Flq|CP000034.1|gene2423|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|GYRA|RequiresSNPConfirmation	5	0	0	0	0	0	0	0	0	0	0	0	0
Flq|CP001138.1|gene3329|Fluoroquinolones|Fluoroquinolone-resistant_DNA_topoisomerases|PARC|RequiresSNPConfirmation	1	0	0	0	2	0	0	0	0	0	0	0	0
CARD|phgb|JQ340367|19-685|ARO:3003582|PmrA|Cationic_antimicrobial_peptides|Polymyxin_B_resistance_regulator|pmrA	3	0	0	0	0	0	0	0	0	0	0	0	0

Chapter 6

Genomic characterisation of sulphite reducing bacteria isolated from the dairy production chain

6.0 Abstract

Anaerobic sporeformers, specifically spoilage and pathogenic members of the genus *Clostridium*, are a concern for producers of dairy products, and of powdered dairy products in particular. As an alternative to testing for individual species, the traditional, and still current, approach to detecting these sporeformers, including non-spoilage/non-pathogenic species, in dairy products has involved testing for a sulphite reducing phenotype (Sulphite reducing Clostridia (SRCs)) under anaerobic conditions. This phenotype is conserved throughout the genus *Clostridium*. Unfortunately, however, this phenotype is also exhibited by other sulphite reducing bacteria (SRBs), leading to potential for false positives. Here, this risk was borne out in that, in addition to species belonging to *sensu stricto*, *Lachnospiraceae* and Cluster XIV of the Clostridia, several SRBs from industry samples were identified as *Proteus mirabilis* and various *Bacillus/Paenibacillus* sp.. Genome wide comparison of a number of representative SRCs and SRBs was employed to determine phylogenetic relationships, especially among SRCs, and to characterise the genes responsible for the sulphite reducing phenotype. This screen identified two associated operons i.e. *asrABC* in SRCs, and *cysJI* in *Bacillus/Paenibacillus spp.* and *P. mirabilis*. Ultimately, this study highlights the inaccuracy of the industry standard SRC test but highlights the potential to generate an equivalent molecular test designed to detect the genes responsible for this phenotype in clostridia.

6.1 Introduction

Raw milk is populated by a variety of metabolically and taxonomically diverse bacteria, the majority of which are inactivated by commercial pasteurization (Wells-Bennik, Driehuis et al. 2016). While this process reduces the overall bacterial load and diversity of the milk, it selects for thermoduric and, in particular, sporeforming, bacteria. This is notable because sporeforming bacteria, including many anaerobic sporeformers, are present in niches throughout the dairy chain, extending from farm to factory (Wells-Bennik, Driehuis et al. 2016) and are a significant concern for the dairy industry (Doyle, Gleeson et al. 2015). The majority of strictly anaerobic sporeformers of concern to the dairy industry belong to the *Clostridium* genus, specifically to Cluster I and Cluster II, and are also known as the *Clostridium sensu stricto* (McAuley, McMillan et al. 2014, Doyle, Gleeson et al. 2015). From a spoilage perspective, some of these *Clostridium* spp. can cause late-blowing defects in cheese due to butyric acid production (Bassi, Puglisi et al. 2015). *Clostridium tyrobutyricum* is most commonly associated with this defect but *Clostridium sporogenes*, *Clostridium butyricum*, *Clostridium beijerickii* and, to a lesser extent, *Clostridium tertium* may also cause or contribute to this defect include (Bermúdez, González et al. 2016). From a public health perspective, *Clostridium perfringens*, *Clostridium botulinum* and *Clostridium tetani* are of greatest concern due to their toxigenic potential. *C. perfringens* is the most prevalent of these species in foodborne illness, and causes in excess of 1 million incidences of foodborne illness in the United States per annum (Scallan, Hoekstra et al. 2011). Although, only one case of foodborne illness in the United States between 1998 and 2008 was attributed to a dairy related vector (Bennett, Walsh et al. 2013), *C. perfringens* has recently been isolated

throughout the dairy farm environment in Australia, including in raw milk (McAuley, McMillan et al. 2014). It has also been detected in defective cheese in Italy (Bassi, Puglisi et al. 2015) and its presence in powdered infant formula (PIF) has been reported (Barash, Hsia et al. 2010). In the case of *C. botulinum*, while the presence of the pathogen in PIF has been associated with two incidences of infant botulism previously, the causative links were not conclusively established (Barash, Hsia et al. 2010). Regardless, this species remains a concern for dairy producers, particularly for those that produce products for infant consumption, as the infectious dose for botulinum spores in infant botulism is thought to be extremely low (ICMSF 2014) and the reputational damage associated with an outbreak would likely be great. Indeed, the inaccurate reporting of the presence of *C. botulinum* in PIF originating from New Zealand recently resulted in a significant product recall (Doyle and Glass 2013). To our knowledge *C. tetani* has not been associated with any incidences of foodborne illness associated with the consumption of dairy product, nor has it been reported to have been detected in dairy products. Nonetheless it remains of concern to producers because of its ability to produce a neurotoxin.

Because of the toxigenicity of some members of the Clostridia, coupled with the potential of some members of the *sensu stricto* to cause spoilage in dairy products, it is routine to test dairy products for the presence of these sporeformers. The test employed most frequently, primarily for historical reasons, involves the enumeration of sulphite reducing Clostridia (SRC) and relies on the ability of the majority of *Clostridium* spp. of concern to the dairy industry to reduce sulphite to sulphide (Weenk, Van den Brink et al. 1995, Doyle, Gleeson et al. 2015), most frequently through cellular enzymes encoded by the *asrABC* operon involved in dissimilatory sulphite reduction (Czyzewski and Wang 2012). However, other bacteria

referred to as sulphite reducing bacteria (SRBs) may have other genes (*cysJI*) that produce the same phenotype (Standards 2003) and result in false positives (Weenk, Fitzmaurice et al. 1991, Doyle, Gleeson et al. 2015). Indeed, aerobic sporeformers and even Gram negative bacteria have caused such false positive results in the past (Sugiyama 1951, Fischer, Zhu et al. 2012). Ultimately, the distribution of the SRC phenotype throughout the heterogeneous *Clostridium* genus, including many species that were previously considered *Clostridium*, (Ludwig, Schleifer et al. 2009) is not well understood, making the relevance of the SRC assay unclear.

The objectives of this study were to determine the identity of SRCs, and SRBs, isolated from a variety of dairy sources, and to employ comparative genomics to identify genetic features common among SRCs with a view to the identification of conserved loci that could be used for alternative, DNA-based, diagnostic approaches.

6.2 Materials and methods

6.2.1 Isolation and identification of sulphite reducing isolates

Anaerobic sulphite reducing bacteria were isolated from dairy powders, cheese and raw bulk tank milk using standard protocols (Standards 2004). This method includes a heat inactivation step (80^{o}C for 10 minutes) that is intended to eliminate non-sporeforming bacteria. Black colonies were then aseptically picked and grown in pure culture in reinforced *Clostridium* media before DNA was extracted using the Mericon Bacteria plus kit (Qiagen). The 16S rRNA gene was amplified from each isolate using the CO1 and CO2 primers (Simpson, Stanton et al. 2003). This PCR was conducted using the following parameters;

94°C for 5 minute, followed by 30 amplification cycles, each consisting of three 1 minute stages at 94°C, 60°C, and 72°C , with a final extension of 5 minutes at 72°C. Amplified DNA was then purified using the GenElute PCR cleanup kit (Sigma Aldrich, Wexford, Ireland) before Sanger sequencing was carried out (Source Bioscience, Waterford, Ireland). The resulting sequences were than subjected to BLAST analysis (Altschul, Madden et al. 1997) against the NCBI database with a view to determining their identity.

6.2.2 Genome sequencing

Genomic DNA, extracted as described above, was further purified using the Powerclean kit (Mo Bio, Carlsbad, CA). Genomic DNA was then quantified using the Qubit high sensitivity kit (Bioscience, Dublin, Ireland), prepared for sequencing using the Nextera XT library preparation kit (Illumina) and sequenced on the Illumina Miseq platform using paired-end 2×250 base pair reads at the Teagasc Sequencing Centre, Teagasc Food Research Centre, Moorepark. Raw reads were processed and filtered based on quality and quantity and trimmed to 200 bp with a combination of Picardtools (https://github.com/broadinstitute/picard) and SAMtools (Li, Handsaker et al. 2009). Quality was visualised using FastQC (Andrews 2010). Sequences were assembled using IDBA-UD (Peng, Leung et al. 2012), removing all contigs smaller than 500bp.

6.2.3 Annotation, phylogenetic comparison and analysis of core genes of *Clostridium* genus

Assembled contigs from sequenced isolates and genome scaffolds from the NCBI genome repository were annotated using Prokka (Seemann 2014). Global alignment of amino acid sequences was carried out using Phylophlan (Segata, Börnigen et al. 2013). A phylogenetic tree was created from this alignment using FastTree (Price, Dehal et al. 2009). The

phylogenetic tree was then visualised using Graphlan (Asnicar, Weingart et al. 2015). Using the .gff files from Prokka, Roary (Page, Cummins et al. 2015) was used to compare the annotated genes from all SRBs using a BLASTp threshold of 50. In addition, core genes within SRC were also identified using Roary (Page, Cummins et al. 2015) setting a BLASTp threshold of 50 % for both comparisons.

6.2.4 *In-silico* screening for sulphite reducing genes among SRBs

A protein database was created containing all the annotated genomes of the SRBs listed in table S1. For the SRC phenotype, query amino acid sequences for the *A*, *B* and *C* subunits of the *asr* gene cluster from the type *C. butyricum* strain, DSM 10702, were BLASTed against this database (Altschul, Madden et al. 1997). For the non-SRC SRB blastp query searches, the amino acid sequences for the assimilatory sulphite reducing genes *cysI* and *cysJ* from *B. licheniformis* were selected as this was the most frequently isolated *Bacillus* SRB in the surveillance.

6.2.5 Analysis of amino acid sequence homology in *asrABC* and *cysIJ*

The sample sequences for BLASTp hit for each gene were retrieved from the BLASTp searches and converted into fasta format and aligned using MUSCLE (Edgar 2004) for visual inspection of conservation. Aligned sequences from each gene were visualised using Jalview (Waterhouse, Procter et al. 2009). The amino acid sequences of the A, B and C subunits of the *asr* operon were examined for the presence of conserved functional domains. Furthermore, the *cysI* and *cysJ* genes were also analysed for conserved amino acid domains. The structure of these proteins was also modelled using Phyre2 (Kelley, Mezulis et al. 2015).

6.3 Results and Discussion

6.3.1 Identification of SRBs in dairy products

In order to better understand the prevalence and identity of SRBs in the Irish dairy chain, 101 positive SRB isolates were identified by sequencing of their corresponding full length 16S rRNA gene amplicons. 77 isolates were identified as clostridia (SRCs), 19 were *Bacillus sp.*, 3 isolates were *Proteus mirabilis* and 2 *Paenibacillus sp.* (Table 1). It was thus apparent that the SRBs present in the dairy chain were relatively heterogeneous, with the proportion of non-clostridia being particularly notable in light of the purpose of the assay i.e. to detect SRCs. The basis for positive phenotypes was anticipated to reflect the presence of *asrABC* operons (i.e. those associated with *Clostridium* spp. (Czyzewski and Wang 2012)), and *cysJI* operons (i.e. those previously found in *P. mirabilis*, *Bacillus/Paenibacillus* and other genera (Guillouard, Auger et al. 2002, Turnbull and Surette 2008)).

Among the 101 isolates, the pathogens detected were *C. perfringens* and *C. tetani*. Although not a pathogen, the presence of *C. sporogenes* is notable in that it can be difficult to distinguish between *C. sporogenes* and *C. botulinum* because of the significant genomic synteny shared between the two species. *C. sporogenes* may also contribute to gas defects in continental style cheeses (Bermúdez, González et al. 2016). The presence of *C. tyrobutyricum*, *C. beijerinckii* and *C. tertium* is notable as these species have previously been associated with late blowing defects in cheese (Cocolin, Innocente et al. 2004, Bermúdez, González et al. 2016). Other clostridia detected were *C. amygdalinum*, *C. bifermentans*, *C. algidcarnis*, *C. aminovelerium*, *C. peptidoveorans*, *C. sartagoforme*, *C. thiosulfatireducens*, *C. cochlearium* and *C. celecrescens*. Of these, *C. bifermentans* has previously been associated with a paediatric infection (Brook 1995) and both it and *C. cochlearium* have previously been

Table 4: SRB detected in the surveillance of raw milk and dairy products.

ID	Source
[Clostridium] amygdalinum	BTM
[Clostridium] amygdalinum	BTM
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	BTM
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry

[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
[Clostridium] bifermentans	Industry
Clostridium tyrobutyricum	BTM
Clostridium algidicarnis	Industry
Clostridium aminovalericum	BTM
Clostridium cochlearium	Industry
Clostridium magnum	Industry
Clostridium pasteurianum/Clostridium beijerinckii	Industry
Clostridium pasteurianum/Clostridium beijerinckii	Industry
Clostridium peptidivorans	BTM
Clostridium peptidovorans DPC 7177	BTM
Clostridium perfringens	BTM
Clostridium perfringens	BTM
Clostridium saratogoforme	BTM
Clostridium saratogoforme	BTM
Clostridium sartagoforme	Industry
Clostridium sartagoforme	BTM
Clostridium sartagoforme	BTM
Clostridium sporogenes	BTM
Clostridium sporogenes	Industry

Clostridium sporogenes	BTM
Clostridium sporogenes	BTM
Clostridium tertium	Industry
Clostridium tertium	Industry
Clostridium tetani	BTM
Clostridium thiosulfatireducens DPC 7172	Industry
Clostridium thiosulfatireducens DPC 7172	Industry
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	Industry
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	Industry
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM

Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	Industry
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridium tyrobutyricum	BTM
Clostridum sporogenes	BTM
Clostridium celecrecens	Industry
Paenibacillus	Industry
Paenibacillus thermophilus	Industry
Proteus mirabalis	BTM
Proteus mirabilis	BTM
Proteus mirabilis	BTM
B cereus HKG	Industry
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus	Industry
Bacillus	Industry
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM

Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM
Bacillus licheniformis	BTM

isolated from powdered infant formula (Barash, Hsia et al. 2010) as well as dairy farm effluent (Gupta and Brightwell 2017), the latter observation potentially highlighting a source of these microbes in the dairy chain. To our knowledge, the presence of *C. amygdalinum*, *C. algidcarnis*, *C. aminovelerium*, *C. peptidoveorans*, *C. sartagoforme*, *C. thiosulfatireducens* and *C. celecrescens* has not previously been reported in dairy sources.

Among the non-clostridia were 16 *Bacillus licheniformis* and 1 *B. cereus* strains. These are spoilage and pathogenic species, respectively, that have been associated with dairy foods (McHugh, Feehily et al. 2017), bulk tank milk (BTM) (Miller, Kent et al. 2015, Sadiq, Li et al. 2016). Finally, three SRB isolates were identified as *P. mirabilis*. The detection of this Gram negative bacterium was unusual as it would be expected that *Proteus* would be inactivated by the heat-treatment step in the assay. Regardless, it is notable that *P. mirabilis* (Kawabata 1980) and *B. licheniformis* (Harmon, Kautter et al. 1971, Weenk, Van den Brink et al. 1995, Fischer, Zhu et al. 2012), though not *B. cereus*, have previously been found to cause false positive results in a SRC assay**.** Two *Paenibacillus spp.*, including *Paenibacillus thermophilus*, were isolated in this screen. *P. thermophilus* has frequently been isolated from raw milk and processed dairy products previously (Ivy, Ranieri et al. 2012).

6.3.2 SRB *in-silico* genome characterisation

In-silico genome characterisation was utilised to further investigate SRB taxonomy, and associated sulphite reducing genes. This analysis included genome sequences that were representative of the species detected in the dairy products and were already available on the NCBI database, as well as sequences corresponding to other cluster 1 *Clostridium*, including *C. botulinum* group and other known sulphite reducing *Clostridium spp.* and members of the *sensu stricto*, *Paenibacillus lactis* (due to the non-availability of a *P.*

thermophilus genome sequence) and *Salmonella enterica typhimurium* LT2, as this species has both an *asrABC* gene cluster for dissimilatory sulphite reduction and a cys*IJ* operon for assimilatory sulphite reduction. Six additional *Clostridium* strains isolated from this study, i.e., *Clostridium aminovelericum* DPC 7173, *Clostridium thiosulfatireducens* DPC 7172, *Clostridium cochlerium* DPC 7174, *Clostridium tertium* DPC 7175, *Clostridium amygdalinum* DPC 7176 and *Clostridium peptidovorans* DPC 7177, were selected for genome sequencing due to the absence of publically available genome sequences at the time of analysis. A list of all the genomes used for this analysis and a summary of the assembly statistics can be found in Table S1.

After sequence assembly and annotation, global genome alignment was carried out using Phylophlan and Roary. Phylophlan uses 300 marker genes common to all bacteria, while Roary uses all the annotated genes of each genome and looks it the presence and absence of these genes, based on a predetermined BLASTp threshold value. The Phylophlan tree (Fig.1) highlights the phylogenetic diversity which exists across bacteria with the SRB phenotype. The division between species that use the *cysJI* operon to reduce sulphite to sulphide (*Bacillus* spp., *P. lactis*, *P. mirabilis* and *S. enterica*) and those that use the *asr* operon is apparent. A similar functional separation, i.e. consistent with the presence or absence of *asrABC*, is observed in the gene presence absence Multidimensional scaling (MDS) plot generated from the Roary results (Fig.2). As noted, *S. enterica* has both *asrABC* and *cysJI* operons. In the Phylophlan tree, the newly sequenced *C. amygdalinum* and *C. aminovelericum* genomes cluster closely with that of *C. celecrescens*. In addition, *C. thiosulfatireducens* shows relatedness to *P. bifermentans* and *C. difficile*. These six species form a distinct branch which is distant from the rest of the *Clostridium* spp. Indeed, some of

species falling within this subgroup have been recently reclassified; for instance the bacterium formally known as *Clostridium difficile*

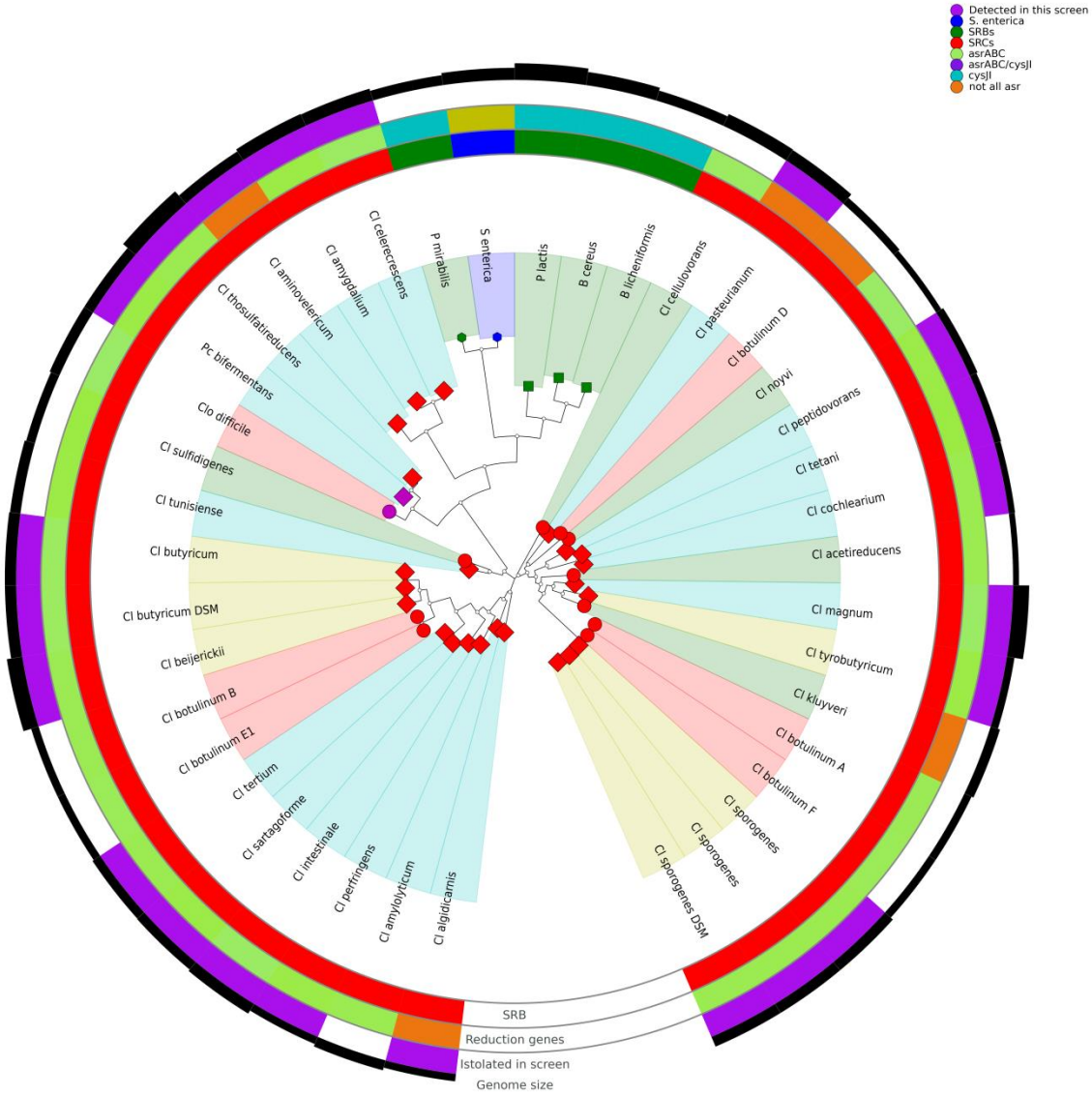


Figure 2: The phylophlan tree is annotated to highlight SRBs isolated during this surveillance.

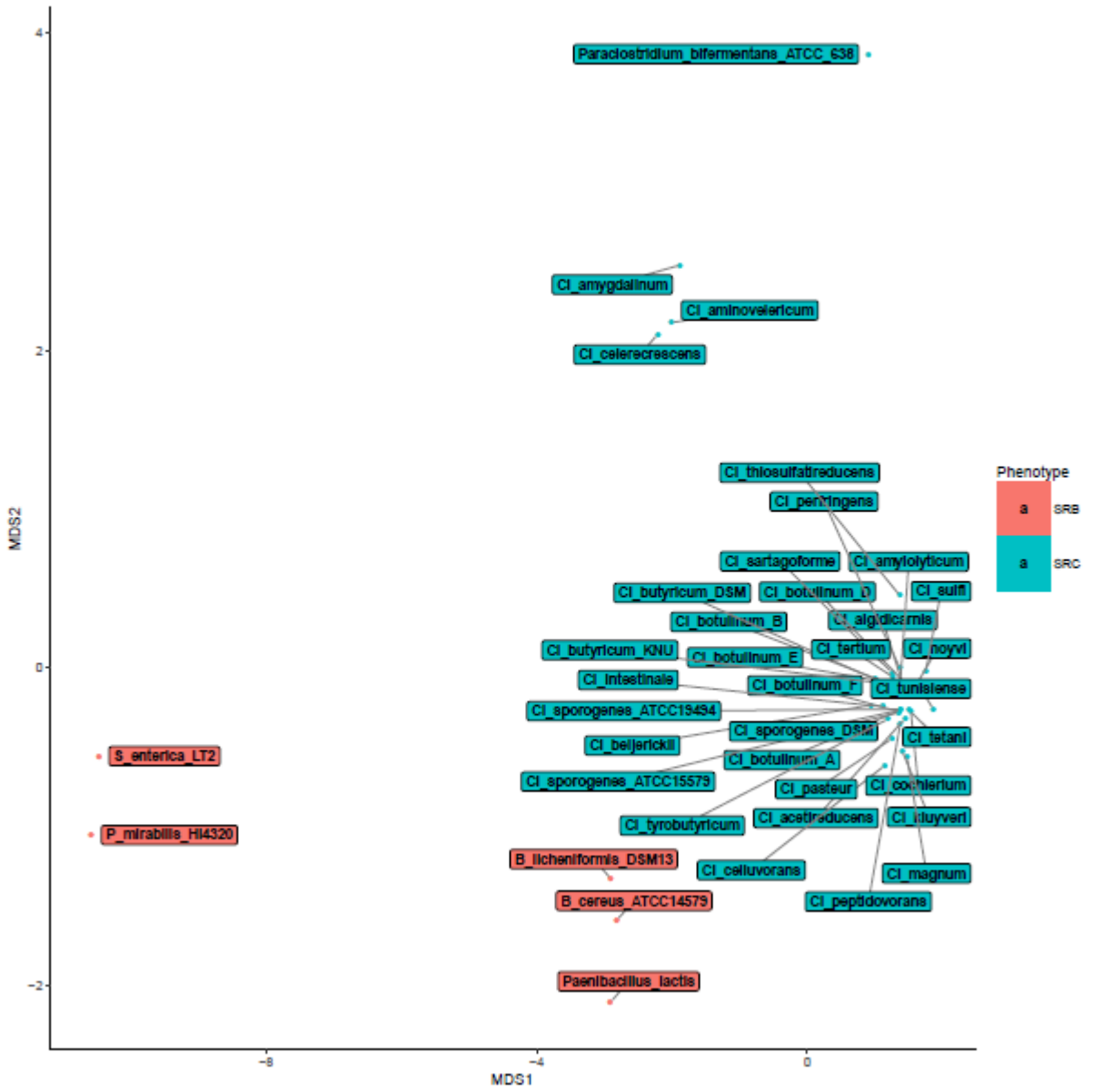


Figure 2: Bray Curtis PCoA depicting the dissimilarity of all SRB genomes; this PCoA is faceted based on the phylogeny.

is now designated as *Clostridioides difficile* (Lawson, Citron et al. 2016). The separation of *C. amygdalinum* and *C. aminovelericum* from the *sensu stricto* can be seen in Fig.2. This is expected as they are members of *Clostridiu*m Clusters XIV and III respectively. It is also evident that the genomes of *C. tunisiense* and *C. sulfidigenes* form a distinct clade separate from the rest of the *sensu stricto Clostridium spp*. Neither of these bacteria were isolated during the present study but were included in the analysis as they are known to reduce sulphite (Thabet, Fardeau et al. 2004, Sallam and Steinbüchel 2009). The distinct clustering of these strains was not anticipated and warrants future investigation. This separation of these two species in Fig.1 suggests that they belong to a distinct subcluster within the *sensu stricto*.

Both the Phylophlan tree and the MDS plot highlight the diversity of sulphite reducing microbes of interest to the dairy industry. While a great number of bacteria can reduce sulphite to sulphite via different pathways (Dahl and Friedrich, 2008), it would appear from these analyses that it is only bacteria which utilise the *asrABC* or the *cysIJ* operons which give a positive test for the SRC assay employed by dairy producers. More specifically, the clostridia that utilise the *asrABC* sulphite reduction pathway are of most concern as they include pathogenic and spoilage-associated bacteria belonging to the genus *Clostridium*. These results highlight the heterogeneity that exists within the Clostridia. While this has already been shown from the context of the 16S rRNA gene sequence (Wiegel, Tanner et al. 2006), genome-wide heterogeneity has until now has not been examined for this Order of bacteria. Although many *Clostridium spp.* have been reclassified and placed with new or existing genera (Lawson, Citron et al. 2016), there is still an issue with *Clostridium* nomenclature. For instance, *C. aminovelericum* and *C. celecrescens* belong to Cluster III of

the clostridia (Wiegel, Tanner et al. 2006), while *C. amygdalinum* belongs to clostridia Cluster XIV based on its phylogeny and high GC content (Parshina, Kleerebezem et al. 2003). The SRC phenotype is distributed across this heterogeneous group of bacteria.

6.3.3 Sulphite reducing protein relatedness in dairy-associated SRBs

While the previous section examined the phylogeny of the SRC and SRB groups, this section details the residue identity associated with the proteins responsible for these phenotypes. Annotated complete sulphite reducing gene clusters for all the SRBs in the database are presented in Fig.S1(B). Fig. 3 (A) depicts the BLASTp bit-score results for AsrABC queries for all of the genomes in the constructed SRB database; the bit score is used to highlight proteins that are similar. The AsrA protein sequence is present at a high degree of homology in the majority of *Clostridium* spp.. Furthermore, there were no BLASTp hits for the AsrA query for *C. acetireducens*, *C. algicarnis*, *C. aminovelericum*, *C. botulinum* D, *C. kluyveri*, *C. noyvi* and *C. pasteurianum*. The AsrB protein sequence was present in the all of the *Clostridium* genomes in the database (Fig.S2). However, levels of homology found in *C. aetireducens*, *C. algicarnis*, *C. aminovelericum*, *C. botulinum* D, *C. kluyveri*, *C. noyvi* and *C. pasteurianum* for this query were much lower than that within other *Clostridiu*m genomes. Similarly, for the predicted AsrC protein, the bit-scores for this query were again low for *C. algicarnis*, *C. aminovelericum*, *C. botulinum* D, *C. kluyveri* and *C. noyvi* and no corresponding gene was found in the *C. acetireducens* and *C. pasteurianum* genomes. Furthermore, this sulphite reducing operon is present at a high degree of homology in *sensu stricto Clostridium* spp. Interestingly, *C. acetireducens, C. kluyveri*, *C. algicarnis, C. noyvi* and *C. pasteurianum,* which are members of the *sensu stricto,* did not have the full operon based

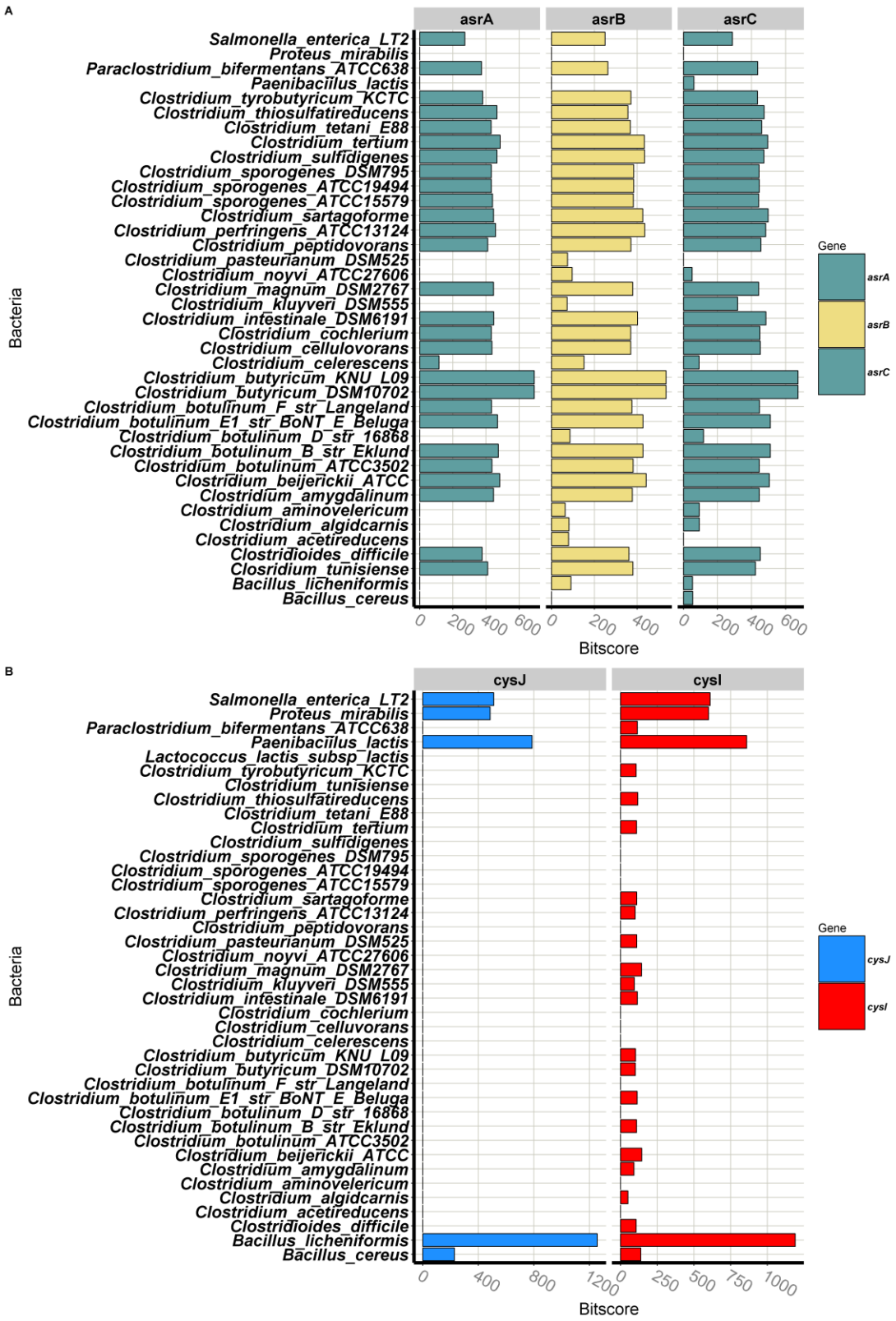


Figure 3: (A) Bar plot depicting bit-scores of BLASTp query hits for asrABC from *C. butyricum* DSM. (B) Bar plot depicting bit-scores of BLASTp query hits for cysJI from *B. licheniformis*.

on these results and highlights that not all *asrABC* genes are necessary to confer the SRC phenotype.

The bar plot in Fig. 3(B) shows the BLASTp results for the CysJI queries for all genomes in the constructed SRB database. For the dissimilatory sulphite reducing pathway involving *cysJI*, the BLASTp bit-scores indicate that, among the SRB genomes placed in the constructed database, the cysJ gene is only present in *B. licheniformis*, *B. cereus*, *P. lactis*, *P. mirabilis* and *S. enterica* Fig.3(B). For the BLASTp with the CysI query, again the highest homology is shared with *B. licheniformis*, *P. lactis*, *P. mirabilis*, *S. enteria* and to a lesser extent *B. cereus.* The presence of this dissimilatory sulphite reductase gene cluster in these species is consistent with what is reported in the literature (Huang and Barrett 1991).

The results from the BLASTp queries of the sulphite reducing genes of *Clostridium* prompted further examination of conserved amino acid domains within AsrABC. Conserved domains could act as targets for a nucleic acid-based detection assay for SRCs as an alternative to the non-specific agar-based approach. It was observed that the proteins AsrA and C contain regions with conserved cysteine motifs. These 4Fe-4S clusters have been observed in AsrA and C in *Salmonella* previously (Huang and Barrett 1991). They have 4 conserved cysteine residues, with a proline toward the C terminus end of the domain. Amino acid sequence alignments can be seen in (Fig. S3). The AsrC protein also contains a siroheme binding site which is annotated in indigo (Fig.S5). AsrB is involved in nucleotide binding (Ostrowski, Barber et al. 1989) (Fig.S4). These alignments show the conservation in the functional regions of these proteins. While similar functional domains might exist in other sulphite reducing bacteria using alternative pathways to the AsrABC mediated reduction, the conserved proline appears to be a unique feature in the Asr 4Fe-4S clusters. To verify that

these conserved domains do not exist in other dairy associated SRBs, we examined the dissimilatory sulphite reducing genes in other SRBs from this analysis. The protein encoded by the *cysJI* operon was also examined for conserved functional domains. The alignment for the alpha-subunit CysJ is shown in (Fig.S6). With conserved YSI and LY motifs observed in the ferredoxin binding domain annotated in indigo in (Fig.S6) and conserved residues in the flavodoxin like domain annotated in cyan. The beta-subunit CysI contains a similar 4Fe-4S cluster to that in AsrA and C (Fig.S7). This sulphite binding cluster does not contain the conserved proline which is a feature of Asr 4Fe-4S clusters. This shows that differences exist not only in the proteins used by these SRBs to reduce sulphite but also in their functional domains.

6.4 Conclusion

Here, the extent to which the agar-based SRC assay fails to distinguish between SRCs and SRBs that are facultative anaerobes was the focus of an extensive investigation. It is apparent that there is a need for a more rapid assay with increased discriminatory power to distinguish between SRCs and the wider group of SRB. Our genome-wide phylogenetic comparison of the dairy-associated SRB phenotype has shown the diversity that exists within this group of microbes. In addition to the noted distribution of this phenotype across Gram positive and negative bacteria, this phenotype is observed throughout the Order Clostridia, with isolates from the *sensu stricto*, *Lachnospiraceae* and Cluster XIV of the Clostridia all producing this phenotype. Furthermore, we have carried out a genomic

characterisation of the SRBs of interest to the dairy industry, with specific focus on Clostridia. This has highlighted the heterogeneity that exists within species that display the SRC phenotype. The wider SRB phenotype can be divided into two further phenotypes based on each isolate's phylogeny and the pathway (AsrABC or CysJI) they utilise to produce the sulphite reducing phenotype. While AsrABC-mediated sulphite reduction has been studied in *S. enterica* and *C. difficile*, it has not been previously examined in the context of SRC phenotype in the dairy industry. Here, we have carried out an *in-silico* screen for the genes of this operon in dairy-associated SRBs and have provided more clarity to what defines a SRC is on the basis of the presence or absence of the *asrABC* operon.

6.5 References

Altschul, S. F., et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic acids research **25**(17): 3389-3402.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Asnicar, F., et al. (2015). "Compact graphical representation of phylogenetic data and metadata with GraPhlAn." PeerJ **3**: e1029.

Barash, J. R., et al. (2010). "Presence of soil-dwelling clostridia in commercial powdered infant formulas." The Journal of pediatrics **156**(3): 402-408.

Bassi, D., et al. (2015). "Understanding the bacterial communities of hard cheese with blowing defect." Food microbiology **52**: 106-118.

Bennett, S. D., et al. (2013). "Foodborne disease outbreaks caused by Bacillus cereus, Clostridium perfringens, and Staphylococcus aureus—United States, 1998–2008." Clinical Infectious Diseases **57**(3): 425-433.

Bermúdez, J., et al. (2016). "Seasonal occurrence and molecular diversity of clostridia species spores along cheesemaking streams of 5 commercial dairy plants." Journal of Dairy Science **99**(5): 3358-3366.

Brook, I. (1995). "Clostridial infection in children." Journal of medical microbiology **42**(2): 78-82.

Cocolin, L., et al. (2004). "The late blowing in cheese: a new molecular approach based on PCR and DGGE to study the microbial ecology of the alteration process." Int J Food Microbiol **90**(1): 83-91.

Czyzewski, B. K. and D.-N. Wang (2012). "Identification and characterization of a bacterial hydrosulphide ion channel." Nature **483**(7390): 494-497.

Doyle, C. J., et al. (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." Int J Food Microbiol **197**: 77-87.

Doyle, M. E. and K. Glass (2013). "Spores of Clostridium botulinum in Dried Dairy Products."

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic acids research **32**(5): 1792-1797.

Fischer, M., et al. (2012). "Culture Media for the Detection and Enumeration of Clostridia in Food." Handbook of Culture Media for Food and Water Microbiology: 66.

Guillouard, I., et al. (2002). "Identification of Bacillus subtilis CysL, a regulator of the cysJI operon, which encodes sulfite reductase." J Bacteriol **184**(17): 4681-4689.

Gupta, T. B. and G. Brightwell (2017). "Farm level survey of spore-forming bacteria on four dairy farms in the Waikato region of New Zealand." MicrobiologyOpen.

Harmon, S. M., et al. (1971). "Improved medium for enumeration of Clostridium perfringens." Applied microbiology **22**(4): 688-692.

Huang, C. J. and E. L. Barrett (1991). "Sequence analysis and expression of the Salmonella typhimurium asr operon encoding production of hydrogen sulfide from sulfite." J Bacteriol **173**(4): 1544-1553.

ICMSF (2014). Usefulness of testing for Clostridium botulinum in powdered infant

formula and dairy-based ingredients for infant formula. ICMSF.

Ivy, R. A., et al. (2012). "Identification and characterization of psychrotolerant sporeformers associated with fluid milk production and processing." Applied and environmental microbiology **78**(6): 1853-1864.

Kawabata, N. (1980). "Studies on the sulfite reduction test for clostridia." Microbiology and immunology **24**(4): 271-279.

Kelley, L. A., et al. (2015). "The Phyre2 web portal for protein modeling, prediction and analysis." Nature protocols **10**(6): 845-858.

Lawson, P. A., et al. (2016). "Reclassification of Clostridium difficile as Clostridioides difficile (Hall and O'Toole 1935) Prévot 1938." Anaerobe **40**: 95-99.

Li, H., et al. (2009). "The sequence alignment/map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Ludwig, W., et al. (2009). Revised road map to the phylumFirmicutes. Bergey's Manual® of Systematic Bacteriology, Springer**:** 1-13.

McAuley, C. M., et al. (2014). "Prevalence and characterization of foodborne pathogens from Australian dairy farm environments." Journal of Dairy Science.

McHugh, A. J., et al. (2017). "Detection and Enumeration of Spore-Forming Bacteria in Powdered Dairy Products." Frontiers in Microbiology **8**.

Miller, R. A., et al. (2015). "Spore populations among bulk tank raw milk and dairy powders are significantly different." Journal of Dairy Science **98**(12): 8492-8504.

Ostrowski, J., et al. (1989). "Characterization of the flavoprotein moieties of NADPH-sulfite reductase from Salmonella typhimurium and Escherichia coli. Physicochemical and catalytic properties, amino acid sequence deduced from DNA sequence of cysJ, and comparison with NADPH-cytochrome P-450 reductase." Journal of Biological Chemistry **264**(27): 15796-15808.

Page, A. J., et al. (2015). "Roary: rapid large-scale prokaryote pan genome analysis." Bioinformatics **31**(22): 3691-3693.

Parshina, S. N., et al. (2003). "Soehngenia saccharolytica gen. nov., sp. nov. and Clostridium amygdalinum sp. nov., two novel anaerobic, benzaldehyde-converting bacteria." International journal of systematic and evolutionary microbiology **53**(6): 1791-1799.

Peng, Y., et al. (2012). "IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth." Bioinformatics **28**(11): 1420-1428.

Price, M. N., et al. (2009). "FastTree: computing large minimum evolution trees with profiles instead of a distance matrix." Molecular biology and evolution **26**(7): 1641-1650.

Sadiq, F. A., et al. (2016). "The heat resistance and spoilage potential of aerobic mesophilic and thermophilic spore forming bacteria isolated from Chinese milk powders." Int J Food Microbiol **238**: 193-201.

Sallam, A. and A. Steinbüchel (2009). "Clostridium sulfidigenes sp. nov., a mesophilic, proteolytic, thiosulfate-and sulfur-reducing bacterium isolated from pond sediment." International journal of systematic and evolutionary microbiology **59**(7): 1661-1665.

Scallan, E., et al. (2011). "Foodborne illness acquired in the United States—major pathogens." Emerg Infect Dis **17**(1).

Seemann, T. (2014). "Prokka: rapid prokaryotic genome annotation." Bioinformatics: btu153.

Segata, N., et al. (2013). "PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes." Nature communications **4**.

Simpson, P., et al. (2003). "Genomic diversity and relatedness of bifidobacteria isolated from a porcine cecum." J Bacteriol **185**(8): 2571-2581.

Standards, I. O. o. (2003). Microbiology of food and animal feeding stuffs - Horizontal method for the enumeration of sulfite-reducing bacteria growing under anaerobic conditions. ISO 2003.

Standards, I. O. o. (2004). Microbiology of food and animal feeding stuffs -- Horizontal method for the enumeration of Clostridium perfringens -- Colony-count technique. ISO 2004.

Sugiyama, H. (1951). "Studies on factors affecting the heat resistance of spores of Clostridium botulinum." J Bacteriol **62**(1): 81.

Thabet, O. B. D., et al. (2004). "Clostridium tunisiense sp. nov., a new proteolytic, sulfur-reducing bacterium isolated from an olive mill wastewater contaminated by phosphogypse." Anaerobe **10**(3): 185-190.

Turnbull, A. L. and M. G. Surette (2008). "L-Cysteine is required for induced antibiotic resistance in actively swarming Salmonella enterica serovar Typhimurium." Microbiology **154**(11): 3410-3419.

Waterhouse, A. M., et al. (2009). "Jalview Version 2—a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191.

Weenk, G., et al. (1991). "Selective enumeration of spores of Clostridium species in dried foods." Journal of applied microbiology **70**(2): 135-143.

Weenk, G., et al. (1995). "Modified methods for the enumeration of spores of mesophilic Clostridium species in dried foods." Int J Food Microbiol **27**(2): 185-200.

Wells-Bennik, M. H., et al. (2016). "Prospects for improved control of dairy-relevant sporeformers using-omics technologies." Current Opinion in Food Science **10**: 38-44.

Wiegel, J., et al. (2006). An introduction to the family Clostridiaceae. The prokaryotes, Springer**:** 654-678.

Supplementary material

Table S1: Genomes in constructed SRB database.

Species	Strain	GC %	Size Mb
Bacillus cereus	ATCC 14579	35.3	5.42
Bacillus licheniformis	DSM 13	46.2	4.22
Clostridium acetireducens	DSM 10703	26.7	2.4
Clostridium algidcarnis	B3	3.06	30.3
Clostridium beijerickii	ATCC 35702	30	6.49
Clostridium botulinum	B str Eklund	27.47	3.8
Clostridium botulinum	ATCC 3502	28.19	3.9
Clostridium botulinum	D str 16868	28.17	3.08
Clostridium botulinum	E1 str BoNT E Beluga	27	3.99
Clostridium botulinum	F str Langeland	28.29	4.01
Clostridium butyricum	DSM 10702	28.5	4.59
Clostridium butyricum	KNU L09	32	3.82
Clostridium celerescens	AAU1	27.9	3.98
Clostridium celluvorans	743B	31.2	5.26
Clostridium intestinale	DSM 6191	30.1	4.6
Clostridium kluyveri	DSM 555	32.02	4.02
Clostridium magnum	DSM 2767	32.1	6.63
Clostridium noyvi	ATCC 27606	27.57	2.61
Clostridium pasteurianum	DSM 525	29.9	4.35
Clostridium perfringens	ATCC 13124	28.4	3.26
Clostridium sartagoforme	AAU1	27.9	3.98
Clostridium sporogenes	ATCC 15579	28	4.1

Clostridium sporogenes	ATCC 19494	27.9	4.06
Clostridium sporogenes	DSM 795	28	4.14
Clostridium sulfidigenes	113A c1	30	3.72
Clostridium tetani	E88	28.8	2.87
Clostridium tunisiense	TJ C661	31.2	4.31
Clostridium tyrobutyricum	KCTC	31.1	3.13
Clostridioides difficile	630	29.1	4.2
Paenibacillus lactis	154	51.8	6.8
Proteus mirabilis	HI4320	38.8	4.09
Paraclostridium bifermentans	ATCC 638	28.4	3.6
Clostridium amygdalinum	DPC 7176	40	5.3
Clostridium peptidovorans	DPC 7177	33	3.4
Clostridium aminovelericum	DPC 7173	35	4.4
Clostridium thiosulfatireducens	DPC 7172	28.3	6.7
Clostridium cochlearium	DPC 7174	29.2	2.6
Clostridium tertium	DPC 7175	28.9	3.6
Salmonella enterica LT2	LT2	52.2	4.95

Table S2: Core genes of SRC phenotype identified by Roary analysis.

Gene	Annotation	No. isolates
accB	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	33
lon1	Lon protease 1	33
clpC	Negative regulator of genetic competence ClpC/MecB	33
mraY	Phospho-N-acetylmuramoyl-pentapeptide-transferase	33
rpe	Ribulose-phosphate 3-epimerase	33
clpP	ATP-dependent Clp protease proteolytic subunit	33
fmt	Methionyl-tRNA formyltransferase	33
pnp	Polyribonucleotide nucleotidyltransferase	33
trmD	tRNA (guanine-N(1)-)-methyltransferase	33
murB	UDP-N-acetylenolpyruvoylglucosamine reductase	33
group_2420	Radical SAM superfamily protein	33
prkC	Serine/threonine-protein kinase PrkC	33
mazG	Nucleoside triphosphate pyrophosphohydrolase	33
fabD	Malonyl CoA-acyl carrier protein transacylase	33
fabG_2	3-oxoacyl-[acyl-carrier-protein] reductase FabG	33
fabF	3-oxoacyl-[acyl-carrier-protein] synthase 2	33
accC	Biotin carboxylase	33
uppS	Ditrans,polycis-undecaprenyl-diphosphate synthase ((2E,6E)-farnesyl-diphosphate specific)	33
ispG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	33
uvrC	UvrABC system protein C	33
clpX	ATP-dependent Clp protease ATP-binding subunit ClpX	33
prfB	Peptide chain release factor 2	33

rlmN	putative dual-specificity RNA methyltransferase RlmN	33
miaA	tRNA dimethylallyltransferase	33
pgsA	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	33
rimO	Ribosomal protein S12 methylthiotransferase RimO	33
group_2719	dihydrolipoamide dehydrogenase	33
dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase	33
group_2821	glmZ(sRNA)-inactivating NTPase	33
ispH	hypothetical protein	33
dxs_1	1-deoxy-D-xylulose-5-phosphate synthase	33
group_2882	hypothetical protein	33
yrrK	Putative Holliday junction resolvase	33
dnaK_1	Chaperone protein DnaK	33
rpsT	30S ribosomal protein S20	33
gpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	33
topA	DNA topoisomerase 1	33
mnmG	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	33
rsmG	Ribosomal RNA small subunit methyltransferase G	33
rsmH	Ribosomal RNA small subunit methyltransferase H	33
engB	putative GTP-binding protein EngB	33
group_2983	Nucleoid-associated protein	33
yvyD	Putative sigma-54 modulation protein	33
tepA	Translocation-enhancing protein TepA	33

mdeA	methionine gamma-lyase	33
group_3095	alanine racemase	33
rpsR	30S ribosomal protein S18	33
hpt_2	Hypoxanthine-guanine phosphoribosyltransferase	33
tpiA	Triosephosphate isomerase	33
rex	Redox-sensing transcriptional repressor Rex	33
frr	Ribosome-recycling factor	33
group_3184	R3H domain protein	33
tig	Trigger factor	33
recR	Recombination protein RecR	33
prfA	Peptide chain release factor 1	33
greA_1	Transcription elongation factor GreA	33
rpsP	30S ribosomal protein S16	33
sigF_2	RNA polymerase sigma-F factor	33
ssbB	Single-stranded DNA-binding protein SsbB	33
group_3473	hypothetical protein	33
tsf	Elongation factor Ts	33
dnaA	Chromosomal replication initiator protein DnaA	33
sigE	RNA polymerase sigma-E factor precursor	33
pyrB	Aspartate carbamoyltransferase catalytic chain	33
rpoB	DNA-directed RNA polymerase subunit beta	33
group_3678	hypothetical protein	33
rpsO	30S ribosomal protein S15	33
ffh	Signal recognition particle protein	33
lepA	Elongation factor 4	33

fabZ	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	33
rpmF	50S ribosomal protein L32	33
sigF_1	RNA polymerase sigma-F factor	33
rplK	50S ribosomal protein L11	33
group_4134	hypothetical protein	33
glyQS	Glycine--tRNA ligase	33
group_4434	hypothetical protein	33
group_4475	hypothetical protein	33
rpmB	50S ribosomal protein L28	33
def	Peptide deformylase	33
ftsH_2	ATP-dependent zinc metalloprotease FtsH	33
accD	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	33

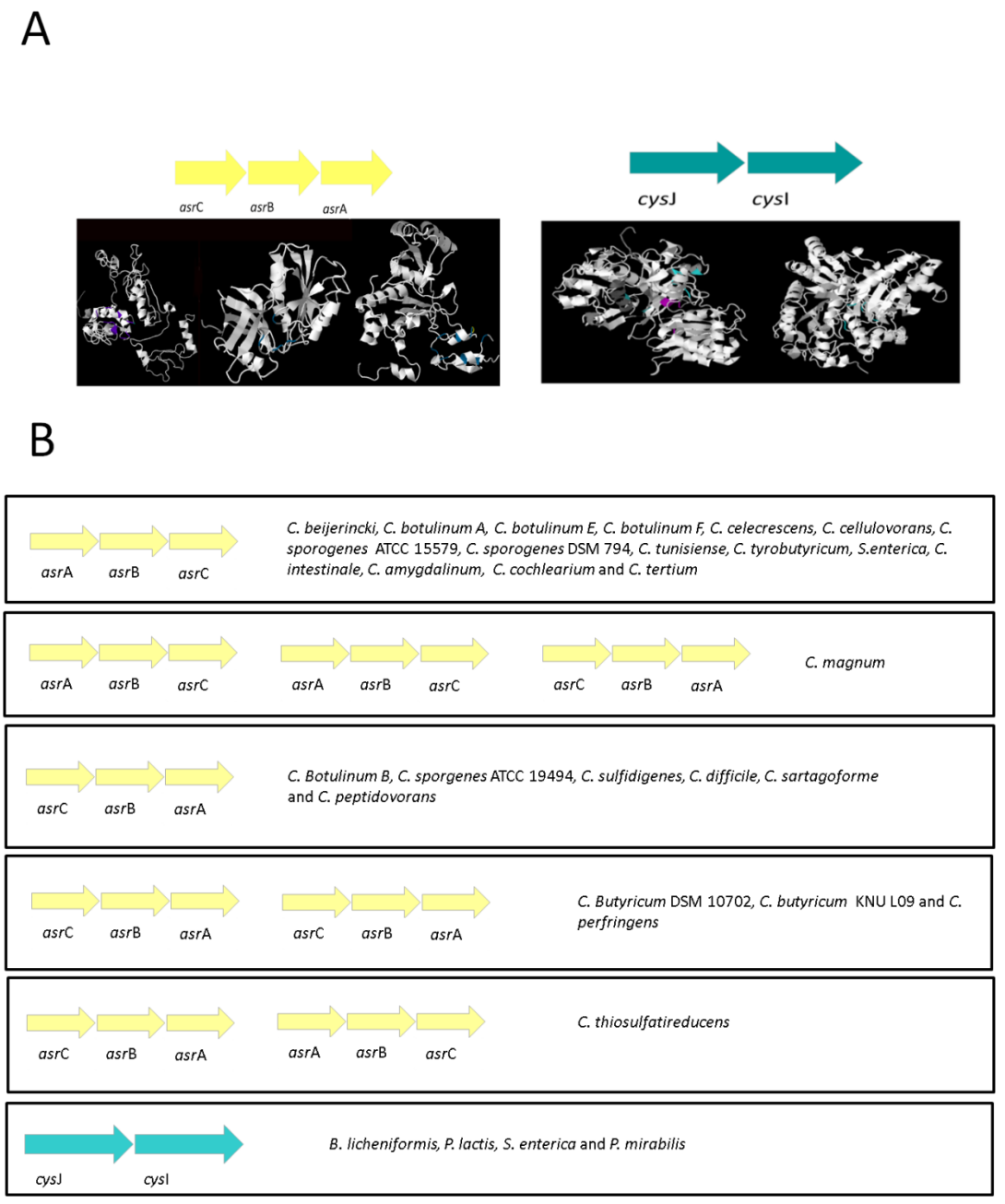


Fig.S1: (A) *asrABC* and *cysJI* sulphite reducing operons, with tertiary structures of proteins, (B) annotated complete sulphite reducing gene clusters in SRBs.

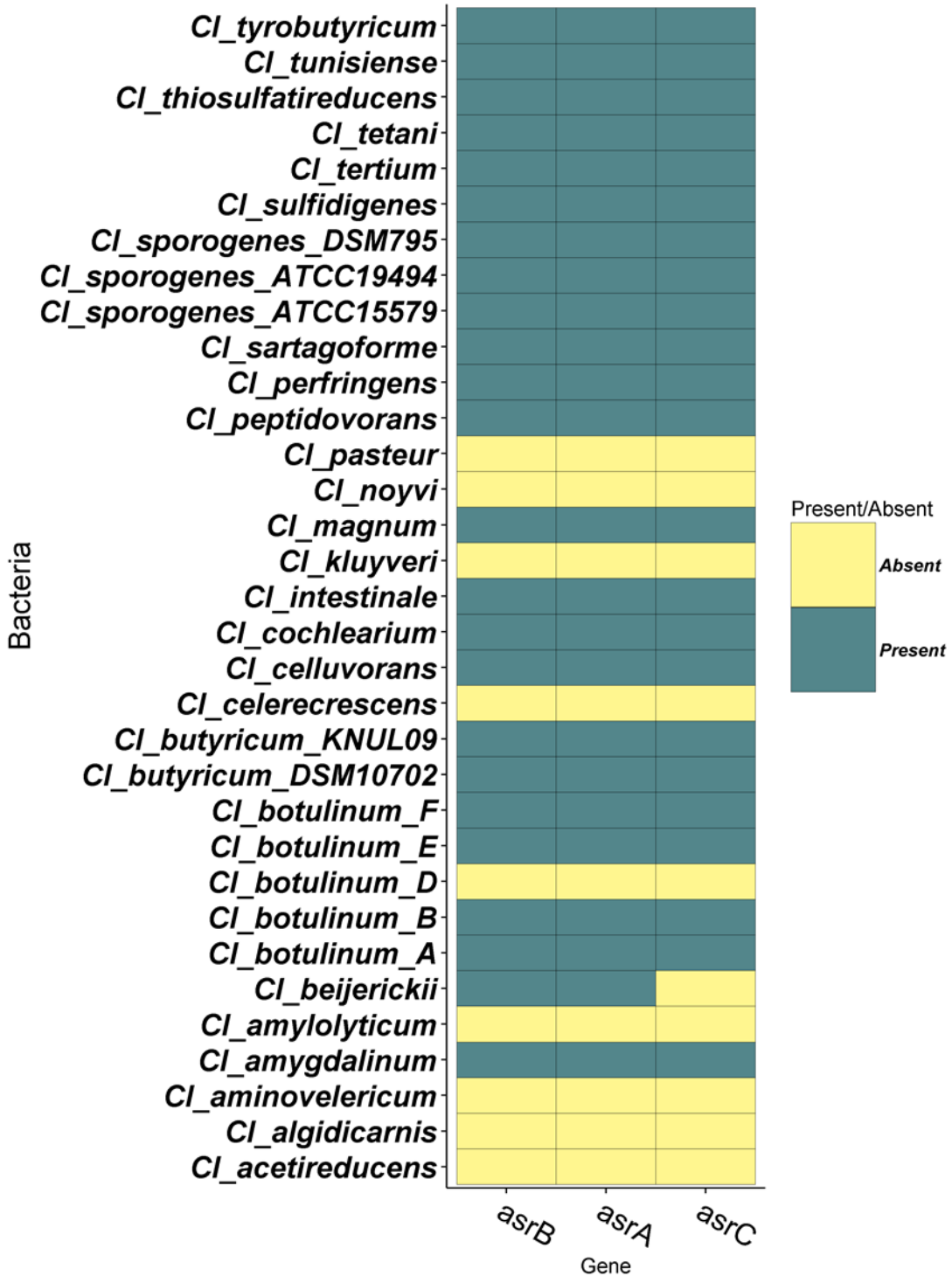


Fig.S2: Roary results for sulphite reducing genes.

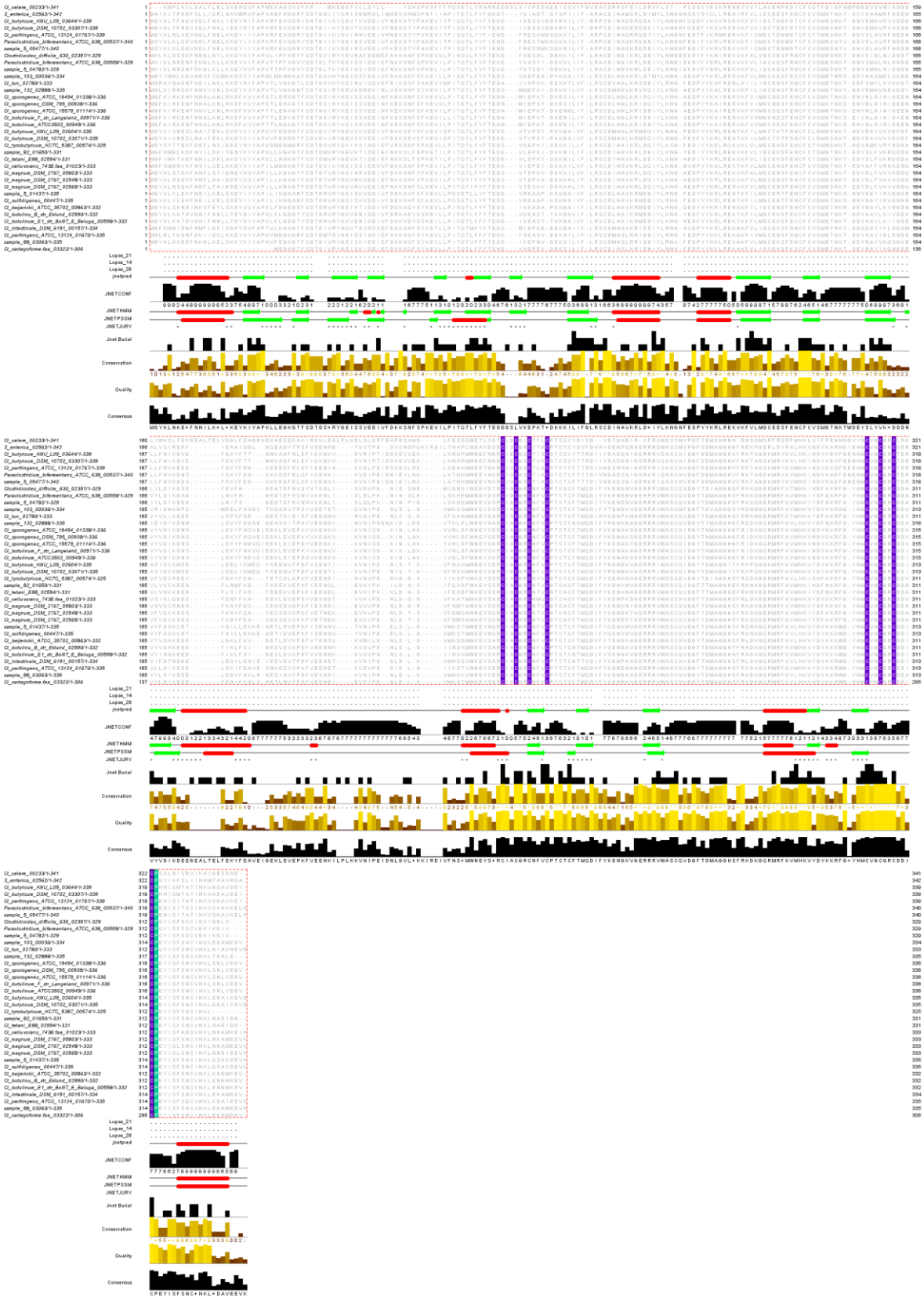


Fig.S3: Jalview alignment of AsrA gene based on BLASTp results; the conserved cysteine and proline residues are highlighted.

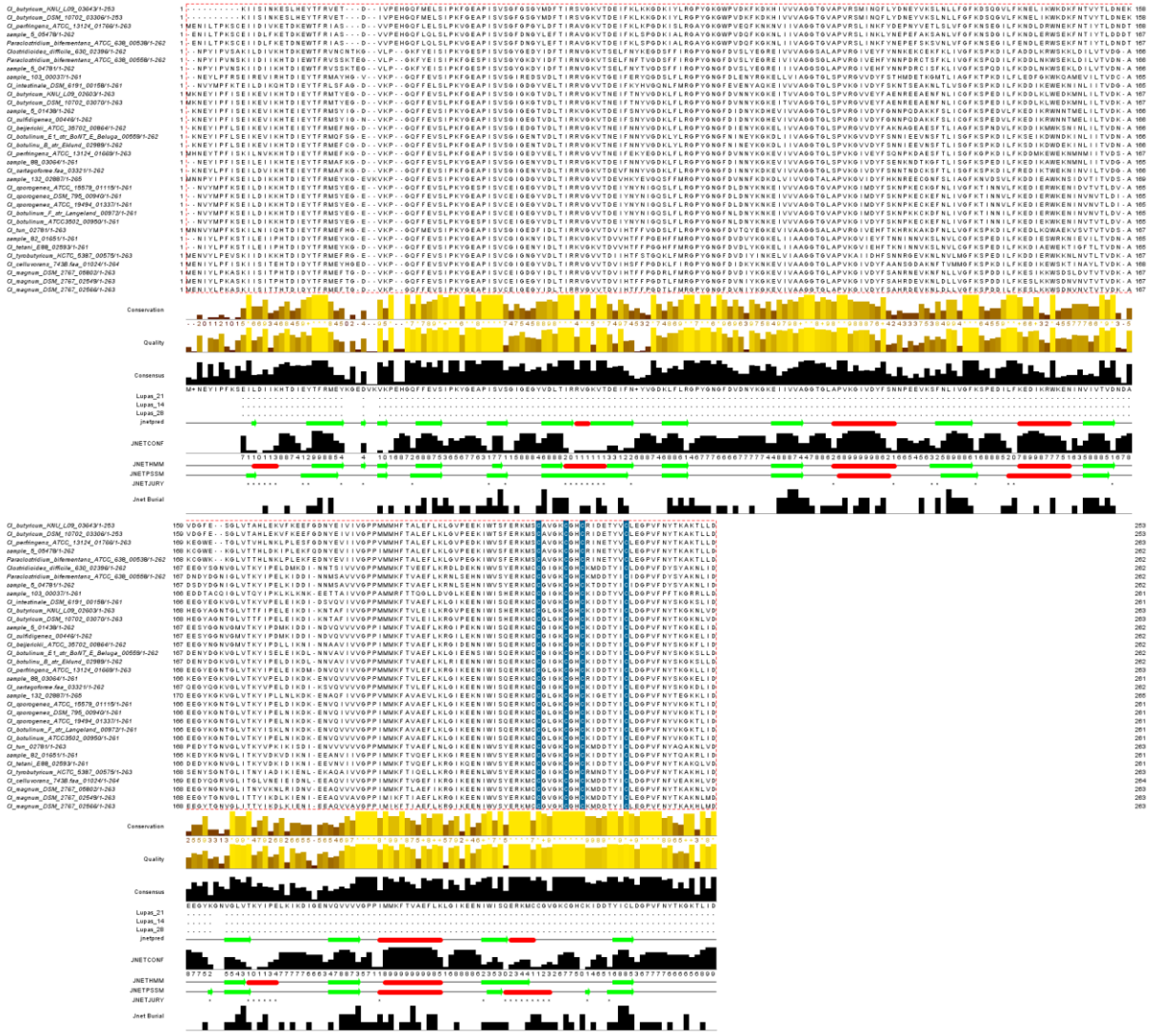


Fig.S4: Jalview alignment of AsrB based on BLASTp results; conserved cysteine residues from this domain are highlighted in blue in the jalview alignments.

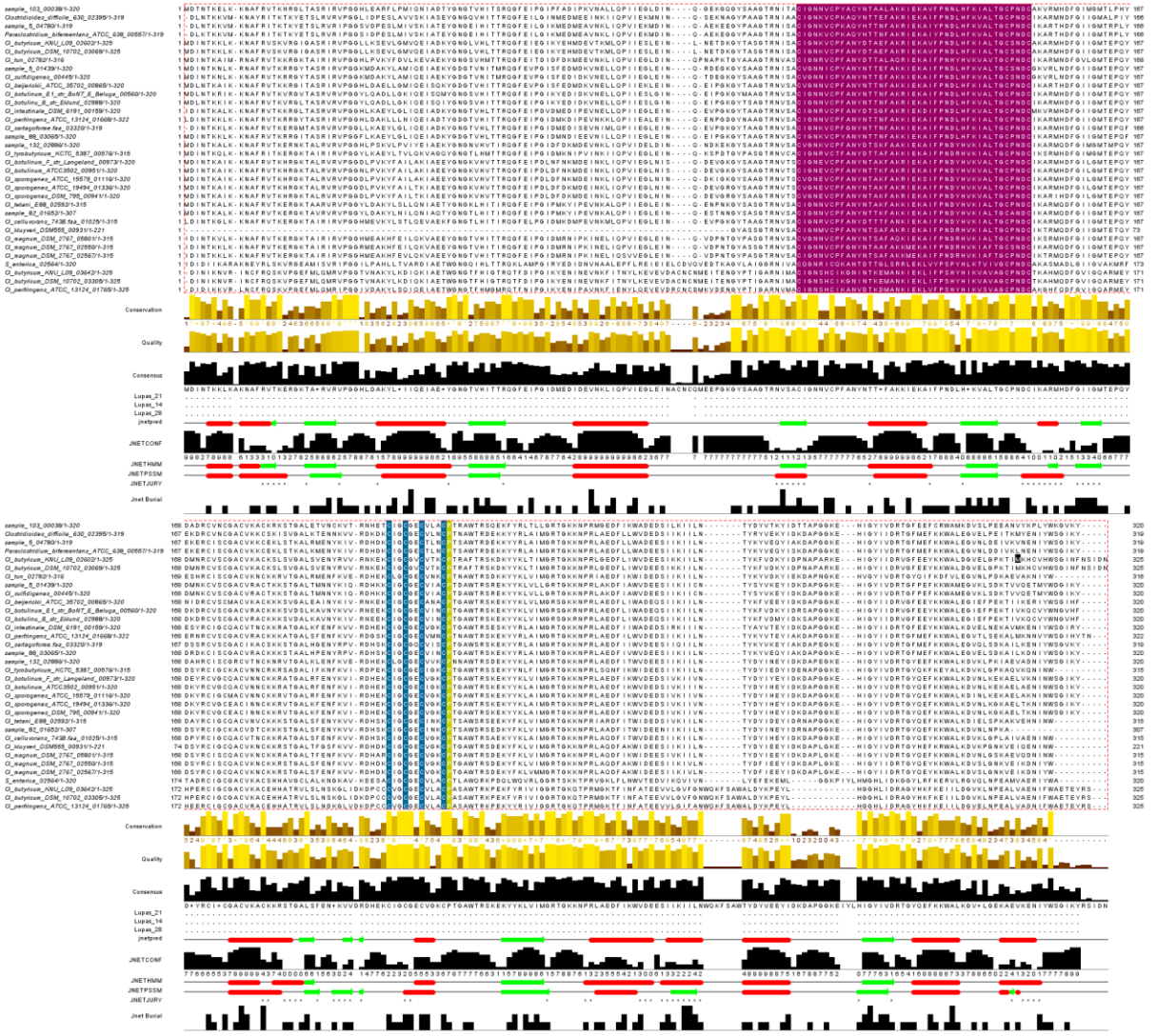


Fig.S5: Jalview alignment of AsrC based on BLASTp results; the conserved siroheme binding site and conserved cysteine and proline residues are highlighted.

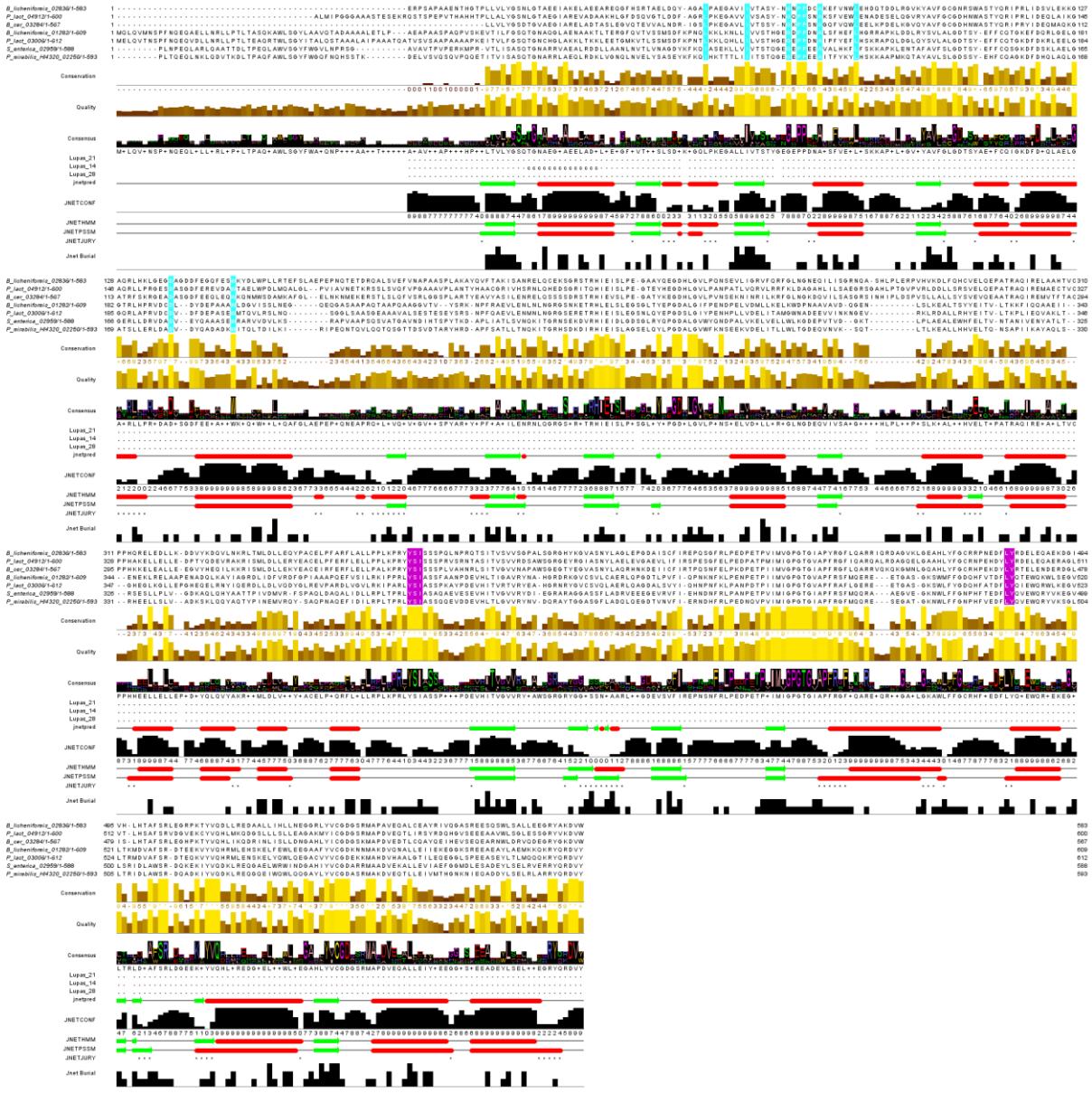


Fig.S6: Jalview alignment of CysJ based on BLASTp results; conserved amino acids in the mononucleotide binding domains, highlighted in the light blue.

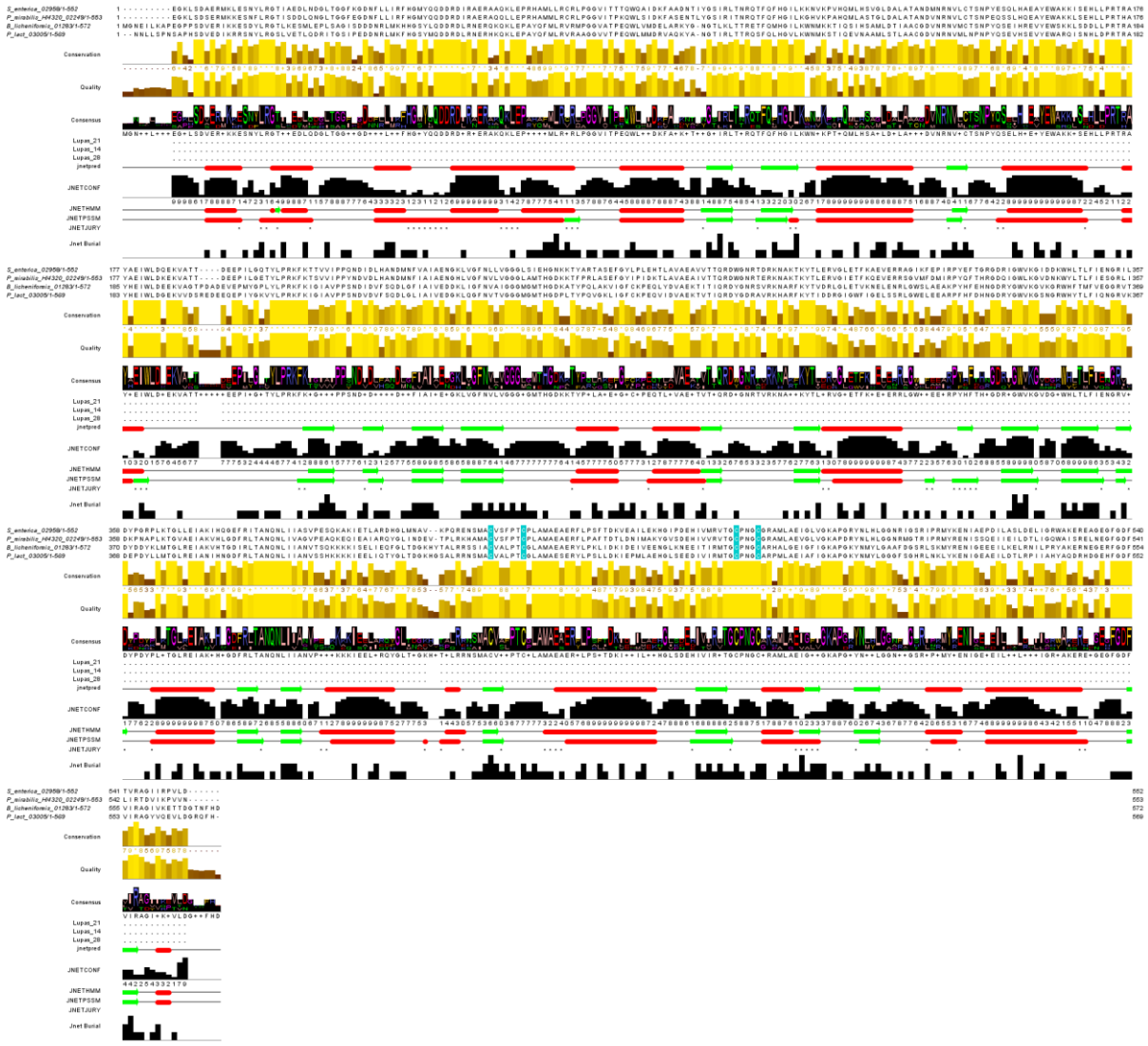


Fig.S7: Jalview alignment of CysI based on BLASTp results; conserved cysteine residues in this amino acid sequence are highlighted.

General discussion

7.0 General discussion

As outlined in Chapter 1, food production chains are home to thriving ecosystems of microorganisms. More specifically to this thesis, the dairy production chain contains numerous niches that may harbour microbes; these niches are present both on farm and in dairy production facilities. It is evident from culture-independent surveillance of both the natural (Gilbert, Jansson et al. 2014, Sunagawa, Coelho et al. 2015) and built environments (Kembel, Meadow et al. 2014) that we are only now gaining a true insight into the plethora of microbes which inhabit these spaces. While culture-based approaches are effective at detecting the presence of potentially pathogenic or spoilage associated microbes, they are generally targeted and require specific media and growth conditions. Indeed, for example, the recent culture independent investigation has identified *Thermus* as the causative agent of the cheese defect known as “pinking” (Quigley, O’Sullivan et al. 2016), a fact that is notable because *Thermus* is not readily detected by any standard industrially applied assay.

Having highlighted this limitation with respect to the traditional methods employed for bacterial detection in the food industry, Chapter 2 summarizes some additional issues associated with a common culture-based assay that is frequently applied in the dairy industry to detect the presence of presumptive spore-formers. More specifically, the sulphite reducing Clostridia (SRC) assay is used to detect indicator organisms. This assay targets bacteria (Clostridia) based on the “sulphite reducing” phenotype that they display under anaerobiosis (Doyle, Gleeson et al. 2015). The issue with this approach is that the functional capacity to reduce sulphite under these conditions is not limited to just one group of taxonomically distinct microorganism. Furthermore, the pathway by which different Orders of distinct bacteria such as Clostridia and Bacilli reduce sulphite is also

different, yet may produce the same phenotype. In addition to exploring the prevalence of anaerobic spore formers in dairy products and issues with detecting the presence of these microbes in milk and dairy products, Chapter 2 also reviews possible sources of SRCs and discusses how the presence of SRCs in dairy products can be better controlled. It focuses on the presence of toxigenic species of the genus *Clostridium*, specifically *Clostridium botulinum* and *Clostridium perfringens* and it discusses how the implementation of good farm management practices and production practices may aid in the reduction of SRCs in milk and dairy products (Gleeson, O'Connell et al. 2013).

The control of bacterial populations in raw milk through the use of different management practices is also explored in Chapters 3 and 4. In these experimental chapters, high-throughput sequencing (HTS) was applied in combination with qPCR to determine how different practices and conditions affect the raw milk microbiota. In Chapter 3, HTS and qPCR were employed to assess the impacts of seasonality, storage temperature and duration on the raw milk microbiota. Results from this study found that storage temperature and duration had little impact on the microbiota, with only increases in the proportions of Gammaproteobacteria such as *Pseudomonas* and *Acinetobacter* being detected in 6^{o}C samples at day 5. This result is of note as these microbes are commonly associated with milk spoilage (Raats, Offek et al. 2011). Indeed, *Pseudomonas* is able to produce heat stable lipases (Sørhaug and Stepaniak 1997), which remain active after pasteurisation, making controlling its presence in raw milk particularity important. Seasonality (lactation stage) had the most influence on the microbiota, with OTUs belonging to Actinobacteria being found in significantly higher proportions in late-lactation compared to mid lactation. A similar observation has also been recently reported in milk from

transgenic goats (McInnis, Kalanetra et al. 2015). Additionally, a significant increase in the presence of *Clostridium* was also noted in late-lactation. This group of bacteria were also found to be higher in a culture based screen of the same samples (O'Connell, Ruegg et al. 2016). One of the limitations of this study was that it targeted DNA and not RNA, this meant that changes in bacterial gene expression in response to different refrigeration temperatures could not be measured. In future, with the application of metatranscriptomic analysis, it may be possible to better understand the dynamics of the raw milk microbiota and how bacteria in the milk adapt to different storage temperatures.

Chapter 4 focused on the influence that seasonal housing and teat preparation had on the raw milk microbiota. This was achieved by sequencing environmental and raw milk samples in parallel and then by inferring possible sources of contamination by utilising the SourceTracker algorithm (Knights, Kuczynski et al. 2011). This method has previously been applied to residential environments such as kitchens (Flores, Bates et al. 2013) and restrooms (Flores, Bates et al. 2011). Perhaps more relevantly, it has also been utilised to track microbial movement in breweries (Bokulich, Bergsveinson et al. 2015). SourceTracker analysis identified faecal and teat surface bacteria as the key contributors of microbes to raw milk from the environment. In the beta diversity analysis we found a clear separation in raw milk microbiota both in BTM and in samples taken from individual cows. This separation was driven primarily by herd habitat; teat preparation had a limited impact on the microbiota. Upon further investigation gut and skin associated bacteria where found in higher proportions in milk from cows housed indoors, whereas soil and environmental type bacteria were found to be more prevalent in milk when the herd was grazing on pasture. Furthermore, it was observed that there were considerable differences between BTM and

individual cow raw milk samples. This suggests that there is an additional contribution of bacteria from the milking equipment to the BTM microbiota. By extending this type of approach in the future it will be possible to quantify the microbial contribution of the milk equipment to raw milk microbiota, to determine if it harbours potentially pathogenic or spoilage-associate microbes and examine how cleaning regimes may influence the transmission of microbes from this niche into the dairy chain.

Following on from this surveillance of the dairy farm environment, Chapter 5 examines the cheese production microbiome before and after the production of continental style cheese by utilising shotgun metagenomic sequencing. This analysis detected the presence of bacteria and phage in cheese production associated samples. Interestingly, Kraken analysis was able to detect resident lactic acid bacteria (LAB) species on production plant surfaces prior to the commencement of cheese-making. *Lactococcus lactis*, a common LAB, was not used as a starter culture for the specific type of cheese in production during the study but was found to dominate production plant surfaces and presumably became established during previous production runs of different cheese types. Resident LAB bacteria have previously been described in American (Bokulich and Mills 2013) and Italian cheese production plants (Calasso, Ercolini et al. 2015). The observation of this phenomenon in three geographically distinct production plants producing different types of cheeses further strengthens the hypothesis that starter bacteria are resident constituents of the cheese production microbiome. Moreover, strain level analysis of the *Lactobacillus helveticus* and *Streptococcus thermophilus* pangenomes identified strains different to common dairy starters which were present in the cheese production plant and in brine samples. The ability to characterise the distribution of bacteria at strain level in the food production

environment highlights how useful this approach could be to monitor contamination patterns during food production. A similar approach has recently been used to detect pathogenic strains of *Escherichia coli* in nunu, an African fermented dairy beverage (Walsh, Crispie et al. 2017). In our study, the detection of phage is notable as phage result in considerable production losses due to failed fermentations (Mahony and van Sinderen 2015). The capacity for HTS to detect phage and bacteria in production plant samples suggests that this approach could be implemented in an industrial context to monitor and control the food production microbiome.

Finally, Chapter 6 investigates the phylogeny of SRCs by 16S rRNA typing of strains, followed by *in-silico* genome characterisation facilitated by whole genome sequencing. This study identified two groups of phylogenetically and functionally different bacteria which produce a positive result in the SRC assay. The first of these is the target group, the SRC, which utilise the *asrABC* operon to reduce sulphite to sulphide. Members of this group belong to the Clostridia. The second group that produce this phenotype are the sulphite reducing bacteria (SRBs). They are not a specific target group for this assay but are able to reduce sulphite using the *cysJI* operon, producing a false positive result. Members of this group include facultative anaerobic bacilli (*Bacillus licheniformis*, *Bacillus cereus* and *Paenibacillus thermophilus*) and *Proteus mirabilis*. As well as highlighting the inaccuracy of the currently applied agar assay to specifically detect SRCs, protein alignments were used to identify conserved amino acid domains in AsrA and AsrC proteins which could potentially be used as targets for a degenerate PCR-based detection of the corresponding genes.

In conclusion, this thesis shows how HTS can revolutionise our understanding of the microbiome of food and of food production environment. In this thesis, the application of

this technology has enabled us to study the influence of seasonality and management practices on the raw milk microbiota. It has identified teat and faeces as the primary sources of bacterial contamination in raw milk and, finally, this work has demonstrated the usefulness of applying HTS in the context of monitoring bacterial species, strains and even phage in the microbiome of dairy foods and dairy production/processing environments. With the advantages of applying HTS to study the dairy production microbiome highlighted throughout this thesis, it is evident that this technology has the potential to be applied to monitor the microbiome of other food production chains too. There are obstacles to implementing HTS to achieve this, such as the current cost of sequencing, storage and computational power requirements for analysing metagenomic datasets and the necessity to validate/gain accreditation for methods for detecting potentially pathogenic and spoilage bacteria in the food production microbiome. However, it is likely that with concerted efforts from food microbiologists, bioinformaticians, technicians and regulatory bodies this type of approach will become the gold standard in microbial surveillance of food production chains in the not too distant future.

References

Bokulich, N. A., et al. (2015). "Mapping microbial ecosystems and spoilage-gene flow in breweries highlights patterns of contamination and resistance." eLife **4**: e04634.

Bokulich, N. A. and D. A. Mills (2013). "Facility-specific “house” microbiome drives microbial landscapes of artisan cheesemaking plants." Applied and environmental microbiology **79**(17): 5214-5223.

Calasso, M., et al. (2015). "Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant." Food microbiology.

Doyle, C. J., et al. (2015). "Anaerobic sporeformers and their significance with respect to milk and dairy products." Int J Food Microbiol **197**: 77-87.

Flores, G. E., et al. (2013). "Diversity, distribution and sources of bacteria in residential kitchens." Environmental microbiology **15**(2): 588-596.

Flores, G. E., et al. (2011). "Microbial biogeography of public restroom surfaces." PloS one **6**(11): e28132.

Gilbert, J. A., et al. (2014). "The Earth Microbiome project: successes and aspirations." BMC biology **12**(1): 69.

Gleeson, D., et al. (2013). "Review of potential sources and control of thermoduric bacteria in bulk-tank milk." Irish Journal of Agricultural and Food Research **52**(2): 217-227.

Kembel, S. W., et al. (2014). "Architectural design drives the biogeography of indoor bacterial communities." PloS one **9**(1): e87093.

Knights, D., et al. (2011). "Bayesian community-wide culture-independent microbial source tracking." Nature methods **8**(9): 761-763.

Mahony, J. and D. van Sinderen (2015). "Novel strategies to prevent or exploit phages in fermentations, insights from phage–host interactions." Current opinion in biotechnology **32**: 8-13.

McInnis, E. A., et al. (2015). "Analysis of raw goat milk microbiota: Impact of stage of lactation and lysozyme on microbial diversity." Food microbiology **46**: 121-131.

O’Connell, A., et al. (2016). "The effect of storage temperature and duration on the microbial quality of bulk tank milk." Journal of Dairy Science **99**(5): 3367-3374.

Quigley, L., et al. (2016). "Thermus and the Pink Discoloration Defect in Cheese." mSystems **1**(3): e00023-00016.

Raats, D., et al. (2011). "Molecular analysis of bacterial communities in raw cow milk and the impact of refrigeration on its structure and dynamics." Food microbiology **28**(3): 465-471.

Sørhaug, T. and L. Stepaniak (1997). "Psychrotrophs and their enzymes in milk and dairy products: quality aspects." Trends in Food Science & Technology **8**(2): 35-41.

Sunagawa, S., et al. (2015). "Structure and function of the global ocean microbiome." Science **348**(6237): 1261359.

Walsh, A. M., et al. (2017). "Strain-level metagenomic analysis of the fermented dairy beverage nunu highlights potential food safety risks." Applied and environmental microbiology: AEM. 01144-01117.

Acknowledgements

Firstly, a sincere thanks to my supervisors Dr. Paul Cotter and Professor Paul O'Toole. Thank you both for your patience, help and support over the past 4 years. Working with you both has been a pleasure. Thanks also to Teagasc for the Walsh Fellowship stipend and to the School of Microbiology in UCC.

Of course none of this would have been possible without the help and support of my colleagues in the Food Bioscience Department in Teagasc Moorepark. Special thanks to Dr. Peter Skuse, Dr. Dan O'Sullivan, Dr. Siobhan Clarke, Dr. Fiona Fouhy, Dr. Aidan Casey, Dr. Mary Rae, Dr. Fiona Crispie and Dr. Clare Piper. To the office crew, working with you guys everyday was an incredibly enjoyable experience. The jokes and the banter were consistent through the good times and the bad. I definitely have made some friends for life in you guys. Special shout-out to Aaron, Calum, Wiley, James and Amy.

Thanks to my parents Mary and Joe, thanks for providing me with support throughout the past 8 years. Words cannot capture how grateful I am to the both of you. To my brothers thanks for keeping me grounded through the last four years. Finally to Ann, thank you for always being there for me and for supporting me the last two years. Thanks for tolerating my final year mood swings. I couldn't have done it without you.

Thanks again to everybody…..peace out